

## Remarks on the grammaticalization of identificational clefts

Denis Creissels

Université Lumière (Lyon 2)

[denis.creissels@univ-lyon2.fr](mailto:denis.creissels@univ-lyon2.fr)

<http://deniscreissels.fr>

**Abstract.** Identificational clefts dissociate the assertion of the exclusive identification of a participant in an event from the rest of the information about the event. In all languages, this can be achieved by combining equative predication and participant nominalization, but in the evolution of languages, the routinization of such a construction as the usual way of expressing participant focalization may result in its grammaticalization as a specific type of construction. After proposing to reformulate the usual distinction between ‘pseudo-clefts’ and ‘clefts’ as a distinction between ‘plain clefts’ and ‘grammaticalized clefts’, this article discusses successively the relationship between cleft constructions and the notion of subordination, the changes that may convert plain clefts into grammaticalized clefts, the emergence of focus markers from cleft constructions, semantic aspects of the evolution of clefts, and the trend towards monoclausality in the evolution of clefts.

### I. Introduction

Identificational clefts are constructions in which the assertion of the exclusive identification of a participant in an event is dissociated from the rest of the information provided about the event in question, presented as presupposed.

The present paper not only discusses the grammaticalization processes in which identificational clefts are involved (a question already widely investigated in the literature, see for example Heine and Reh 1984 and Harris and Campbell 1995), but also puts forward a new theoretical perspective on the relationship between the two types of constructions commonly termed ‘clefts’ (*It was the interrogation that frightened him most*) and ‘pseudo-clefts’ (*What frightened him most was the interrogation*).

My claim is that the current terminology obscures the nature of the relationship between these two types of constructions. The point is that the current terminology tends to restrict ‘clefts’ to constructions more or less similar to the *it*-clefts of English, which wrongly suggests that this type constitutes a more primitive type of construction (logically speaking) which can be defined independently, whereas the only possible definition of so-called pseudo-clefts must refer to a more primitive notion corresponding to the *it*-clefts of English, from which ‘pseudo-clefts’ inherit some of their properties, but not all.

On the contrary, I argue that so-called ‘pseudo-clefts’ are a type of construction universally identifiable on the basis of a straightforward and unproblematic definition. By contrast, there is no such definition on the basis of which the constructions designated as clefts in

descriptions of individual languages could be identified as ‘clefts’ CROSS-LINGUISTICALLY, and the only possible CROSS-LINGUISTIC definition of so-called clefts is that referring to their historical relationship with so-called pseudo-clefts.

In this article, after discussing the usual distinction between ‘pseudo-clefts’ and ‘clefts’ and proposing to reformulate it as a distinction between ‘plain clefts’ and ‘grammaticalized clefts’ (section 2), I discuss the relationship between cleft constructions and the notion of subordination (section 3), the changes that may convert plain clefts into grammaticalized clefts (section 4), the possibility that focus markers emerge as the result of the evolution of cleft constructions (section 5), semantic aspects of the evolution of clefts (section 6), and the trend towards monoclausality in the evolution of clefts (section 7). Section 8 summarizes the main conclusions.

## 2. Plain clefts and grammaticalized clefts

As recalled above, identificational clefts are constructions in which the assertion of the exclusive identification of a participant in an event is dissociated from the rest of the information provided about the event in question, presented as presupposed.

In the simplest cases, this is achieved by transparently combining the equative predication construction and the participant nominalization construction,<sup>1</sup> without the intervention of any additional (construction-specific) syntactic rule, as for example in (1): in (1a), a participant nominalization (*what you need*) occupies the pre-verbal subject slot in a plain copular clause. In (1b), the participant nominalization (*the one who told me that*) is found in the clause-final afterthought position, but none of these two constructions requires positing specific rules.

(1) English

(1a) *What you need is a good sleep.*

(1b) *It's John, the one who told me that.*

It has long been observed that, in focus constructions that do not necessarily show the other characteristics commonly considered typical of clefts, the focalized phrase is often systematically found in either clause-initial or clause-final position, combined with a marker

---

<sup>1</sup> The term ‘participant nominalization’, often restricted to an operation of derivational morphology, is taken here in a broad sense. By participant nominalization, I mean any expression whose syntactic distribution is similar to that of NPs and whose internal structure designates its referent as participating with a given role in an event encoded as a verbal lexeme, whatever the exact relationship between the internal structure of the expression in question and the corresponding clause. In this understanding of ‘participant nominalization’, this notion includes not only ‘headless’ or ‘free’ relatives, but also relatives with semantically light heads (i.e., nominal or pronominal heads that may be syntactically required but do not really restrict the set of potential referents, since the restriction they introduce is most of the time implied by the lexical meaning of the verb of the relative clause), such as *the person/one* [*who ...*], *the thing* [*that ...*], in French *celui/celle/ce* [*qui ...*], in Spanish *el/la/lo* [*que ...*], etc.

resembling (or identical to) an equative copula (i.e. a verb or grammatical word that marks nouns as expressing identification), as in examples (2) to (5).<sup>2</sup>

(2) Mina (Rémy Bôle-Richard, pers.com.)

(2a) *Kòfì d̀̀ m̀́lú.*  
Kofi eat rice  
'Kofi ate rice.'

(2b) *M̀́lú yé.*  
rice it\_is  
'It's rice.'

(2c) *M̀́lú yé Kòfì d̀̀.*  
rice it\_is Kofi eat  
'Kofi ate RICE', lit. 'It's rice Kofi ate.'

(3) Zarma (Oumarou Yaro 1993: 108-111)

(3a) *M̀́usà d̀̀i Ábdù.*  
Moussa see Abdou  
'Moussa saw Abdou.'

(3b) *Ábdù ǹ́o.*  
Abdou it\_is  
'It's Abdou.'

(3c) *Ábdù ǹ́o M̀́usà d̀̀i.*  
Abdou it\_is Moussa see  
'Moussa saw ABDOU.' lit. 'It's Abdou Moussa saw.'

(4) Caribbean Spanish (Kato & Mito 2016)

(4a) *Perdió el anillo fue Miguel.*  
lost.3SG the ring was.3SG Miguel  
'It's Miguel who lost the ring.'

(4b) *Nosotras llegamos fue cansaditas.*  
we.F arrived.1PL was.3SG tired.PL.F  
'How we arrived was tired.'

---

<sup>2</sup> In the glossing of the examples, the following glosses are used in addition to those listed in the Leipzig Glossing Rules: CL noun class, CONT continuative, CPL completive, D definiteness or specificity marker, ICPL incomplete, LH light head in participant nominalization, LK linker, POSTP postposition, PRET preterite.

(5) Brazilian Portuguese (Kato & Miotto 2016)

(5a) *Me ajudou muito foi o psicólogo.*  
me helped.3SG much was.3SG the psychologist  
'It's the psychologist who helped me much.'

(5b) *O João deu pra Maria foi o livro.*  
the João gave.3SG to Maria was.3SG the book  
'It's the book that João gave to Maria.'

It can also be observed that, even in focus constructions in which the focused constituent does not combine with a focus marker, the non-focused part of the construction may show properties of relative clauses, in particular in the selection of particular verb forms. For example, in the construction of Hausa illustrated in (6b), the use of a focus marker is not obligatory,<sup>3</sup> but the TAM that follows a fronted phrase in focus role must be a relative form.

(6) Hausa (Newman 2000: 188)

(6a) *Yanà bāyan bishiyà.*  
3SG.M.CONT behind tree  
'He is behind the tree.'

(6b) *Bāyan bishiyà yakè.*  
behind tree 3SG.M.RELCONT  
'It's behind the tree he is.'

Such observations suggest a grammaticalization path in which the source construction expresses the exclusive identification of a participant in an event presented as presupposed by simply combining the equative predication construction and the participant nominalization construction, without involving any additional (construction-specific) syntactic rule, as illustrated above in example (1). It is widely assumed that, at some point in the evolution of a given language, the routinization of such a construction as the usual way of expressing participant focalization may result in its grammaticalization as a specific type of construction, and eventually in its conversion into a monoclausal focus construction.

In this grammaticalization path (analyzed among many others by Givon 1979, Heine & Reh 1984: 109-110, 147-182, Harris & Campbell 1995: 152-162), the source construction is unambiguously biclausal, whereas the final outcome may be a focalizing device that can be described as operating within the frame of the syntax of the simple clause, in which all that remains from the source construction is a focus marker, most commonly (but not always, cf. section 5) a reflex of the former equative copula, or the use of special verb forms, initially motivated by the subordinate status of the part of the construction referring to a presupposed event.

---

<sup>3</sup> According to Katharina Hartmann (presentation at the workshop 'Clefts and related focus constructions'), it is not optional either, since the use of the equative copula in focus marker function can be analyzed as indicating an exhaustive interpretation of the focus constituent.

For example, the grammaticalized focus construction of Jóola Fóoñi illustrated by (7a) is different from the construction expressing the same meaning by simply combining the equative predication construction and the participant nominalization construction (7b). In particular, in Jóola Fóoñi, subject nominalization requires the use of a participle (a non-finite verb-form that does not express person agreement), whereas in (7a), the verb agrees in person with the focalized subject. However, the participle in (7a) and the finite verb form in (7b) share the suffixal inflexion characteristic of relative verb forms, which can be viewed as evidence that (7a) historically derives from a construction not necessarily identical to (7b), but which in any case initially involved relativization.

(7) Jóola Fóoñi (pers.doc.)

(7a) *Ínjé i-jee-ñaa bee ka-sanken-a-k.*  
 1SG 1SG-go.ICPL-REL to INFk-talk-D-CLk  
 ‘It’s me who is going to talk.’

(7b) *A-jee-ñaa bee ka-sanken-a-k, ínjé eenom.*  
 (CLa)PTCP-go.ICPL-REL to INFk-talk-D-CLk 1SG it\_is  
 ‘The one who is going to talk, it’s me.’

Crucially, this grammaticalization path involves a gradual process of clausal fusion, and the final stage can be considered to have been reached when nothing in the functioning of the construction requires positing an underlying combination of two clauses.

In formal syntax, the possibility of deriving grammaticalized focus constructions from underlying structures that simply combine equative predication and participant nominalization has long been acknowledged (cf. among many others Akmajian 1970, Gundel 1977), and there is a voluminous literature devoted to the discussion of the technical details of this derivation.

The constructions expressing the exclusive identification of a participant in a presupposed event by simply combining the equative predication construction and the participant nominalization construction are commonly designated as *pseudo-clefts*. This terminological choice is more than questionable, not only for the reasons already mentioned in the introduction, but also because the term *pseudo-cleft* suggests that the constructions in question do not really involve a cleaving process. Quite on the contrary, the process of cleaving affecting the expression of a propositional content and resulting in a separation between the asserted and presupposed parts of the sentence is particularly clear-cut in this type of construction, whose originality is rather that the separation is achieved by simply combining two constructions that exist independently from each other (the equative predication construction and the participant nominalization construction), rather than by using a special construction.

This is the reason why, following Harris and Campbell (1995: 152), the term *cleft* will be used here to include both the constructions commonly called pseudo-clefts and those termed clefts. My proposal is to replace this terminological distinction by a distinction between plain clefts and grammaticalized clefts. In plain clefts (or ‘transparent’ clefts, in the terminology suggested by Van der Wal & Maniacky 2015), the separation between the part of the sentence that refers to a presupposed event (the *content clause* in Harris & Campbell’s terminology)

and that asserting the identification of a participant (the *clefted constituent*) is achieved by simply combining an equative predication construction and a participant nominalization construction, whereas grammaticalized clefts involve construction-specific rules, and may have discursive functions that are not limited to the exclusive identification of a participant in a presupposed event.

An advantage of this definition is that it makes it clear that plain clefts are among the universally available syntactic constructions (Harris & Campbell 1995: 56). To the extent that a language has an equative predication construction and a participant nominalization construction in the sense defined in Footnote 1 (and it seems difficult to imagine a language that would lack these two types of constructions), it is always possible to combine them to make explicit the exclusive identification of a participant in a presupposed event. Consequently, not every language has grammaticalized clefts, but such constructions easily develop as soon as, for some reason, plain clefts become a usual way of expressing focalization.

An interesting consequence of the universal availability of plain clefts is that limitations to the development of (grammaticalized) clefts in the history of languages can only be imagined in terms of competition between different types of focalization strategies. Given the universal availability of the source construction, it would not make sense to imagine restrictions involving other aspects of morphosyntactic typology.

### **3. Clefts as atypical subordinating constructions**

The clauses involved in complex constructions typically refer to distinct (although related) events (or states-of-affairs), and the function of complex constructions is both to encode the semantic nature of the relationship between the events in question, and to hierarchize them from the point of view of communicative dynamism.

What is special about clefts, and makes them atypical as subordinating constructions, is that the clauses combined into a cleft construction cannot be viewed as referring to distinct events. Cleft constructions do not integrate the representation of two events into a complex syntactic construction, they rather decompose the representation of a single event by positing the property of playing a given role in an event whose other participants are presented as already identified, and asserting the identification of the entity having this property. For example, *It was John who asked Mary for a dance* presupposes the existence of an event that can be described as *Someone invited Mary for a dance*, but does not say anything about possible relationships between this event and any other event. What is asserted by the cleft construction is the identity of the person that has the property of ‘having invited Mary for a dance’.

There is nothing original in this analysis of cleft constructions, but keeping it in mind is crucial for a proper understanding of the changes that characterize the grammaticalization of cleft constructions, and of their tendency to evolve towards monoclausal focus constructions. In some sense, everything happens as if the routinization of a cleft construction expressing the exclusive identification of a participant in an event favored changes whose guiding principle is that they reduce the difference between the cleft construction and the plain assertive clause referring to the same event.

#### 4. The change from plain clefts to grammaticalized clefts

The first stage in the grammaticalization of clefts is that a plain cleft construction undergoes changes that make it impossible to describe the construction as a mere instance of equative predication in which the phrase referring to the entity to be identified has the form of a participant nominalization. Alternatively, the specificity of a cleft construction may be reinforced by changes that affect equative predication or relativization but are not reflected in the cleft construction.

##### 4.1. A morphological element obligatory in the equative predication construction may be optional or absent in the cleft construction

This situation is illustrated by Hausa. In Hausa, “in normal equational sentences, the stabilizer [i.e. equative predicator] is obligatory”, and can only be left out in some specific structures (Newman 2000: 164). By contrast, as illustrated in (8), in the cleft construction, the equative predicator is not obligatory (cf. Introduction, Footnote 2).

(8) Hausa (Newman 2000: 188)

(8a) *Audù nē.*

Audu it\_is.M  
'It's Audu.'

(8b) *Audù yā tàfi kàsuwā.*

Audu 3SG.M.CPL go market  
'Audu went to the market.'

(8c) *Audù (nē) ya tàfi kàsuwā.*

Audu it\_is 3SG.M.REL.CPL go market  
'It is Audu who went to the market.'

##### 4.2. A morphological element obligatory in the participant nominalization construction may be optional or absent in the cleft construction

In French, participant nominalization normally requires an overt head (either a noun or the 'light head' *celui/celle/ce* – etymologically a demonstrative pronoun).<sup>4</sup> This element does not occur in the grammaticalized cleft construction illustrated in (9b).

---

<sup>4</sup> In French, participant nominalization without an overt head can still be found in frozen expressions such as proverbs (cf. for example *Qui veut voyager loin ménage sa monture* 'Slow and steady wins the race', lit. Who wants to travel far must take care of their horse'), but is not productive anymore.

(9) French

(9a) *Je connais celui qui a parlé.*  
 I know the\_one who has spoken  
 ‘I know the one who spoke.’

(9b) *C’est Jean qui a parlé.*  
 it\_is John who has spoken  
 ‘It’s John who spoke.’

In Tswana, relative clauses (either noun-modifying or ‘free’ relatives) are obligatorily introduced by a linker expressing gender-number (‘class’) agreement, as in (10a-b). Historically, there can be little doubt that this linker results from the grammaticalization of a demonstrative, since the grammaticalization of demonstratives into relativizers is an extremely common grammaticalization process, especially in Bantu languages, and in Tswana, the whole paradigm of twelve forms expressing gender-number agreement is exactly the same for the demonstrative and the relativizer. Identification is commonly expressed by clauses whose only obligatory elements are the invariable equative predicator *kí* ‘it is’ and a noun phrase expressing identification, as in (10c). The construction may include a second noun phrase designating the entity to be identified, as in (10d), and the position of this second NP may be occupied by a participant nominalization, giving rise to the plain cleft construction illustrated in (10e). However, there is also a grammaticalized variant of the cleft construction in which the linker is omitted, as in (10f), although the content clause maintains the other characteristics of a relative clause.

(10) Tswana (pers.doc.)

(10a) *Ga ke itse monna yo o rogileng Mpho.*  
 ɣà-kí-ítsí mò-ńnà jó ʼó-róɣílè-ń m:pʰó  
 NEG-1SG-know CL1-man CL1.LK CL1-insult.PRF-REL (CL1)Mpho  
 ‘I don’t know the man who insulted Mpho.’

(10b) *Yo o rogileng Mpho o kae?*  
 jó ʼó-róɣílè-ń m:pʰó ʼó-kâ:i  
 CL1.LK CL1-insult.PRF-REL (CL1)Mpho CL1-where  
 ‘Where is the one who insulted Mpho?’

(10c) *Ke Kitso.*  
 kí kî:tsò  
 it\_is Kitso  
 ‘It’s Kitso.’

(10d) *Ke Kitso mong wa mmotorokara.*  
 kí kítsó mò-ń wá-m-mótóròkâ:rà  
 it\_is 1SG CL1-owner CL1.GEN-CL3-car  
 ‘It’s Kitso, the owner of the car.’



(10e) *Ke Kitso yo o rogileng Mpho.*  
 kǐ    'kítsó   'jó    'ú-róχílè-ń    m̀:phó  
 it\_is 1SG   CL1.LK   CL1-insult.PRF-REL   (CL1)Mpho  
 'It's Kitso who insulted Mpho.'

(10f) *Ke Kitso o rogileng Mpho.*  
 kǐ    'kítsó   'ú-róχílè-ń    m̀:phó  
 it\_is 1SG   CL1-insult.PRF-REL   (CL1)Mpho  
 'It's Kitso who insulted Mpho.'

### 4.3. Person agreement in subject clefts

In languages in which verbs express person agreement with their subject, a very common phenomenon characterizing grammaticalized clefts as opposed to plain clefts is that, in subject focalization, the verb in the content clause agrees with the clefted subject, which is incompatible with the description of the content clause as an instance of participant nominalization. In terms of grammaticalization process, the development of a possibility of person agreement in subject clefts results in making the focus construction more similar to the corresponding plain assertive sentence than it originally was, and can therefore be viewed as evidence of a trend towards reanalyzing the construction as monoclausal.

For example, Akmajian (1970) analyzed the competition between the two variants of the *it*-cleft construction of English illustrated in (11), and observed that all the speakers that allow (11b), with person agreement, also allow (11a), but not *vice versa*, which suggests that the option in (11a) should still be considered as somehow more basic.

(11) English (Akmajian 1970)

(11a) *It is me who is responsible.*

(11b) *It is I who am responsible.*

A similar phenomenon can be observed in many languages. For example, in Tswana, in the cleft construction illustrated in (10e-f) above, if the clefted constituent is a 1st or 2nd person pronoun, the verb in the content clause can optionally agree with it in person, giving rise to the three variants of the cleft construction illustrated in (12). The variant illustrated in (12c) is by far the commonest one in the data I collected, but the other two are also widely attested (and the variant illustrated in (12b) is for example the one systematically used in the Tswana translation of the Bible).

(12) Tswana (pers.doc.)

(12a) *Ke nna yo o apeetseng Mpho.*  
 kǐ    ñná   'jó    ú-àpéétsì-ń    m̀:phó  
 it\_is 1SG   CL1.LK   CL1-cook.APPL.PRF-REL   (CL1)Mpho  
 'It's me the one who cooked for Mpho.'

- (12b) *Ke nna yo ke apeetseng Mpho.*  
 kí òná 'jò kí-àpéétsì-ń m̀:phó  
 it\_is 1SG CL1.LK 1SG-cook.APPL.PRF-REL (CL1)Mpho  
 ‘It’s me who cooked for Mpho.’

- (12c) *Ke nna ke apeetseng Mpho.*  
 kí òná kí-àpéétsì-ń m̀:phó  
 it\_is 1SG 1SG-cook.APPL.PRF-REL (CL1)Mpho  
 ‘It’s me who cooked for Mpho.’

Interestingly, some languages provide evidence that agreement of the verb of the content clause with a clefted subject may develop very early in the history of a cleft construction. For example, contrary to modern Romance languages, Latin was certainly not among the languages in which cleft constructions are widely used to express focalization, but as illustrated in (13), subject cleft constructions in which the verb of the content clause agrees in person with the clefted subject are attested as early as in the plays of Plautus: in a plain cleft construction, the content clause in (13) would be expected to occur as *qui te dudum conduxit* ‘(the one) who hired.3SG you just now’.

- (13) Latin (Plautus, *Mercator* – quoted by Dufter 2008)

*Non ego sum qui te dudum conduxi.*  
 not I am who you just\_now hired.1SG  
 ‘I’m not the person that hired you just now.’

#### 4.4. The flagging of the clefted constituent in grammaticalized clefts

In plain cleft constructions, the clefted constituent shows no variation in flagging related to its role in the content clause, since whatever its role in the content clause, its syntactic role in a plain cleft construction is exclusively and invariably that of the phrase expressing identification in an identificational clause. By contrast, an important aspect of the evolution of cleft constructions is that, in grammaticalized clefts, the clefted constituent may be flagged according to its role in the content clause. Here again, this makes the grammaticalized cleft more similar to the corresponding plain assertive clause, and can therefore be viewed as evidence of a trend towards reanalyzing the construction as monoclausal.

Let’s for example examine the cleft construction of Spanish illustrated in (14).

- (14) Colloquial Peninsular Spanish (Zubizarreta 2014)

- (14a) *De la que te hablé fue de María*  
 of LH.SG.F that 2SG.DAT speak.PRET.1SG be.PRET.2SG of María  
 lit. ‘Of whom I spoke to you was of María.’

- (14b) *Con el que hablé es con Pedro.*  
 with LH.SG.M that speak.PRET.1SG be.PRS.2SG with Pedro  
 lit. ‘With whom I spoke is with Pedro.’

The corresponding plain cleft constructions could be something like [*La persona de la que te hablé*] *fue María* ‘The person of whom I spoke to you was María’ and [*La persona con la que hablé*] *es Pedro* ‘The person with whom I spoke is Pedro’, where the square brackets delimit the phrase expressing participant nominalization. In other words, the construction in (14) differs from the mere combination of equative predication and participant nominalization in two respects: the semantically light noun fulfilling the role of head in participant nominalization is omitted, and the preposition flagging the relative pronoun is repeated before the clefted constituent.

Interestingly, French illustrates the next stage of this evolution, i.e. the stage at which the flagging of the clefted constituent constitutes the only manifestation of its role in the content clause. Example (15) illustrates three types of cleft constructions expressing the focalization of a prepositional phrase in oblique function: with flagging of the relative pronoun only (a), with flagging of both the relative pronoun and the clefted constituent (b), and with flagging of the clefted constituent only (c).

(15) French

- (15a) *C’est toi à qui je veux parler.*  
 it-is you to whom I want to\_speak  
 lit. ‘It’s you to whom I want to speak.’

- (15b) *C’est à toi à qui je veux parler.*  
 it-is to you to whom I want to\_speak  
 lit. ‘It’s to you to whom I want to speak.’

- (15c) *C’est à toi que je veux parler.*  
 it-is to you that I want to\_speak  
 lit. ‘It’s to you that I want to speak.’

As discussed by Dufter (2008), cleft constructions of the type illustrated in (15a), with the absence of flagging of the clefted constituent characteristic of plain clefts, were the norm in Old and Middle French (until the 15th century). They are very rare in present-day French. The type illustrated in (15b), with double flagging, was particularly common in 16th-18th century French. It still occurs sporadically, especially with *de*-marked obliques (cf. for example *C’est d’un cancer dont il souffre*, lit. ‘It’s from cancer from which he suffers’). The type illustrated in (15c), with flagging of the clefted constituent only, is clearly the norm in present-day French, which nicely illustrates the correlation between the increase in the frequency of clefts and the trend towards aligning the coding of predicate-argument relationships (indexation and flagging) in the cleft construction with their coding in the corresponding plain assertive clause.

## 5. Changes in cleft constructions and the emergence of focus markers

### 5.1. Marking of focused NPs resulting from the evolution of a cleft construction

In plain cleft constructions, it is the construction itself that compositionally determines the exclusive identification reading, but as a result of the changes that tend to make cleft constructions less transparent and more similar to the corresponding plain assertive clauses, the focus interpretation tends to be linked to morphological elements whose presence was initially motivated by their role either in the equative predication construction or in the participant nominalization construction, and which have been maintained after the construction has changed in such a way that they cannot be analyzed as fulfilling their initial role anymore.

A first possible scenario, in languages in which relativization requires the use of special verb forms, is that relative marking on the verb is maintained after the construction has changed in such a way that it cannot be analyzed as involving relativization anymore, which implies that the special inflection that initially signaled that the verb is the nucleus of a relative clause is reanalyzed as signaling focalization of one of the terms of the construction, as in the focus construction of Jóola Fóoni illustrated in (5b) above.

Another possible scenario is that an element in contact with the clefted constituent in the source construction is maintained after the construction has changed in such a way that the initial motivation for its presence has been lost, hence its reanalysis as a focus marker. As discussed by Heine & Reh (1984) and Harris & Campbell (1995), depending on the details of the construction in individual languages, this reanalysis process may affect elements whose initial status was that of copula, expletive pronoun, or relativizer.

Still another possible origin of focus markers is discussed by Van der Wal & Maniacky (2015) for a group of Bantu languages including Kituba (the vehicular language of the Western part of the Democratic Republic of Congo) and Lingala (the vehicular language of the North-Western part of the Democratic Republic of Congo). The languages in question have subject focus constructions with focus markers that have grammaticalized from (and are still homonymous with) a noun ‘person’. Van der Wal & Maniacky’s analysis is that, initially, the construction was a plain cleft with the noun ‘person’ in the function of light head in participant nominalization (something like ‘John is the person who invited Mary’ for ‘JOHN invited Mary’). In this function, ‘person’ underwent a process of semantic bleaching that made it available for constructions in which the focalized subject is not necessarily human or animate, and the construction changed in such a way that, as argued by Van der Wal & Maniacky, ‘person’ can be analyzed now as a focus marker in a monoclausal focus construction. (16a-b) illustrate the use of *muntu* (originally: ‘person’) in Kituba as a focus marker in combination with non-human nouns, and (16c) illustrates the possibility that *muntu* as a focus marker co-occurs with the noun *muntu*.

(16) Kituba (Van der Wal & Maniacky 2015)

(16a) *Pusu muntu me kudya mbisi.*  
 cat FOC (<person) PRF eat fish  
 ‘It’s the cat that has eaten the fish.’

- (16b) *Inti muntu me kubwa.*  
 tree FOC (<person) PRF fall  
 ‘It’s the tree that has fallen.’
- (16c) *Muntu muntu me yimbila.*  
 person FOC (<person) PRF sing  
 ‘It’s a person who has sung (not a bird).’

## 5.2. Predicative marking of focused NPs in focus constructions that do not seem to originate from a cleft construction

Some languages have focus constructions involving a focus marker homonymous with an identificational predicator, but in which the focused constituent does not occur systematically in clause-initial or clause-final position, as expected in focus constructions resulting from the grammaticalization of plain clefts. The possibility that the evolution of cleft constructions is not necessarily the only source of focus markers homonymous with identificational predicators must therefore be considered.

This is in particular the case with the focus construction of Ivorian Jula (17c-e), in which the focused constituent is focalized *in situ* by the adjunction of a focus marker identical to the identificational predicator.

- (17) Ivorian Jula (pers.doc.)
- (17a) *Fántà kà jégê tóbí lónân yé.*  
 Fanta CPL.TR fish cook visitor for  
 ‘Fanta cooked fish for the visitor.’
- (17b) *Fántà / Jégé / Lónán lò.*  
 Fanta / fish / visitor it\_is  
 ‘It’s Fanta / fish / the visitor.’
- (17c) *Fántà lò kà jégê tóbí lónân yé.*  
 Fanta it\_is CPL.TR fish cook visitor for  
 ‘FANTA cooked fish for the visitor.’
- (17d) *Fántà kà jégé lò tóbí lónân yé.*  
 Fanta CPL.TR fish it\_is cook visitor for  
 ‘Fanta cooked FISH for the visitor.’
- (17e) *Fántà kà jégê tóbí lónán lò yé.*  
 Fanta CPL.TR fish cook visitor it\_is for  
 ‘Fanta cooked fish for THE VISITOR.’

There can be little doubt that the focus marker in this construction results from an extension of the use of the identificational predicator. It is, however, very dubious that a cleft construction played a role in this evolution, for the following reasons:

- In Jula, this strategy is not limited to constituents in a particular position: any constituent of the clause can be focalized *in situ* in this way.
- All Manding languages have a very productive focus construction of the same type as (26c-e) (i.e., without movement of the focused constituent), with however a focus marker distinct from the identificational predicator in most Manding varieties.
- Jula is supposed to have developed relatively recently as an offshoot of Bambara, where the focus marker and the identificational predicator are *dè* and *dòn*, respectively);
- No grammaticalized cleft construction has been described in any of the Manding languages for which documentation is available, and in all Manding languages, plain clefts are very marginal.

It is probably safer to leave the question open, but taking into consideration the fact that Jula is a language variety that has emerged in a context of intensive language contact with non-Mande languages (Senufo and others), one may wonder whether this use of an identificational predicator as a focus marker operating *in situ* really results from a grammaticalization process in the usual sense of this term, or perhaps rather from the kind of phenomena that typically occur in second-language acquisition. It seems reasonable to assume that equative predication is acquired very early by second-language learners, so that at a stage when they have not acquired the focus construction of the target language, the mere adjunction of the identificational predicator to a constituent of a plain assertive clause constitutes a readily available focalization strategy: the construction resulting from this adjunction may be ungrammatical for native speakers of the target language, but it raises no problem in terms of communicative efficiency, since there can be no doubt about the intended meaning.

## 6. Semantic aspects of the grammaticalization of clefts

As far as the use of a cleft construction remains sporadic in a given language, its function does not depart from the emphatic expression of its compositional meaning, i.e. the exclusive identification of a participant in a presupposed event. An increase in the frequency of a cleft construction automatically triggers a decrease in emphasis and a widening of its possible discourse functions, as analyzed among many others by Ball (1994) and Patten (2012) for English, and by Dufter (2008) for French.

(18) illustrates a possible outcome of the evolution of the discourse function of cleft constructions, in which the departure from the original meaning is particularly clear.

(18) French

- *Est-ce que tu as entendu ce bruit?*  
is\_it that you have heard this noise  
'Did you hear this noise?'

- *C'est le voisin qui passe l'aspirateur.*  
 is\_it the neighbor who passes the-vacuum\_cleaner  
 'It's the neighbor who is vacuuming.'

In this example, it is clear from the context that the second sentence includes no overt reference to the presupposed event about which the speaker is reacting (i.e., the noise), and could be paraphrased as 'What's happening is that the neighbor is vacuuming'. What remains of the original meaning of the cleft construction is that the assertion is about a presupposed event, but the whole sentence consists of new information presented as explaining the event in question. The articulation between the clefted constituent and the rest of the construction can be viewed as highlighting the role of a particular participant ('What's happening is that someone is vacuuming, and this person is the neighbor'), but by no means as reflecting an articulation between presupposition and assertion.

Güldemann (2010) observed a similar polyfunctionality for the cleft constructions of the Tuu languages of Southern Africa, and analyzed it in terms of 'entity-centralthetic statements' in the sense of Sasse (1985).

Note that the possibility of using focalization just to mark that the sentence explains the current situation seems to be a common property of grammaticalized focus constructions, independently of the fact that they are recognizable as resulting from the evolution of a cleft construction or not, cf. among others Creissels & Sambou (2013: 422-423) for Mandinka, Robert (2010) for Wolof.

(19) Creissels & Sambou (2013: 423)

- Ni i yè káanàa jé kùrùtù-kàròo lá*  
 if 2SG CPL.TR iguana.D see trousers-sew-D POSTP  
*fěñé-bóndí-dúlàa lè bé à búlù.*  
 tail-take\_out-place FOC COP 3SG in\_the\_sphere\_of  
 lit. 'If you see an iguana sewing trousers, it's a way to take its tail out that it has.'  
 > ... you can conclude that it has a way to take its tail out. (proverb)

(20) Wolof (Robert 2010)

(situation: a person arrives and hears screaming)

- *Lu xeew fi?*  
 what happen here  
 'What is going on here?'
- *Musaa, moo dóor Ndey.*  
 Moussa 3SG.FOC beat Ndey  
 'It's Moussa who has trashed Ndey.'

Example (21) provides further illustrations of this possibility of a radical departure from the original meaning of a cleft construction. It is clear from the context that, in the second part of the sentence in (21a), the presupposition is not 'I brought her something', but rather something like 'I did something to reconcile with her'. In (21b), it is even more obvious that

the presupposition cannot be ‘There is something that she did not raise’, and the only justification that can be imagined for expressing ‘She didn’t do anything’ via a clefted variant of the figurative expression ‘She didn’t raise a finger’ is that clefting makes the assertion more emphatic.

(21) Spanish (Zubizarreta 2014)

(21a) *No me he disculpado, lo que le he traído es un regalo.*  
I didn’t apologize      LH.N that 3SG.DAT have.1SG brought is a gift  
lit. ‘... what I brought her is a present.’ > ‘... what I did is bring her a present’

(21b) *Lo que no alzó fue un dedo.*  
LH.N that not raised was a finger  
lit. ‘What she did not raise was a finger.’ > ‘What she did was not to raise a finger.’

The use of cleft constructions as a pure narrative device constitutes another radical departure from their original meaning. For example, in French narrative texts, it is quite common to find sentences such as (22) in contexts in which it is clear that the cleft construction must not be interpreted literally as specifying the dating of an event (i.e., as an answer to an implicit question ‘When did the event in question happen?’), since the second part of the sentence refers to an event that has not been mentioned yet by the narrator. The function of such clefts is rather to highlight a major moment of the story (‘What happened at that point is that ...’).

(22) French

*Et c’est alors que la voiture tomba en panne.*  
and it-is then that the car fell in breakdown  
lit. ‘And it is then that the car broke down.’ >  
‘What happened then is that the car broke down.’

Turning to terminological issues, I would like to emphasize in this connection that the term of pseudo-cleft, misleadingly used in the literature on cleft constructions for the constructions I propose to designate as plain clefts, would be rather appropriate for constructions like those in (18), (21), and (22), since they are characterized by a radical mismatch between the formal articulation ‘clefted constituent / rest of the construction’, and the semantic articulation ‘presupposition / assertion’.

## **7. The trend towards monoclausality in the evolution of cleft constructions and the focalization-negation interplay**

In the literature on cleft constructions and their evolution, there is a consensus on the fact that cleft constructions develop from uncontroversially biclausal constructions, but tend to change over time into monoclausal focus constructions. There is a voluminous literature devoted to the discussion of the monoclausal or biclausal status of focus constructions in individual languages, based however on language-specific criteria, without any real discussion of a



comparative concept of clausehood that would make it possible to consistently evaluate the monoclausal / biclausal character of focus constructions across languages. Not surprisingly, as observed for example by Zentz (2016) about the analysis of the cleft constructions of Bantu languages, it is very common that different linguists analyzing the same focus constructions take opposite decisions about their mono- or biclausal nature, and the reason is simply that they do not apply the same diagnostics, or disagree about their relative conclusiveness. It is symptomatic that the controversy about the mono- or biclausal nature of the English *it*-cleft is still raging, in spite of the particularly voluminous literature devoted to detailed investigations of all possible aspects of its behavior – cf. among many others Haegeman et al. (2014).

The question of a comparative concept of clausehood has recently been discussed by Haspelmath (2016) in relationship to serial verb constructions (another syntactic phenomenon for the analysis of which the distinction between mono- and biclausal constructions is crucial). Building on Bohnemeyer & al. (2007) and Comrie (1995), he proposes that, cross-linguistically, a crucial property of monoclausal constructions, as opposed to constructions involving two or more clauses, is that they can only have one way to form the negation. The obvious advantage of this criterion in a cross-linguistic perspective is that, among the various criteria that may have been evoked in discussions about mono- vs. biclausality, it is the only one whose universal applicability is unquestionable. In this perspective, the other criteria manipulated in discussions about mono- vs. biclausality can be viewed as more or less typical properties of mono- or biclausal constructions, but by themselves, they are not conclusive.

However, the observation of negation in cleft constructions casts some doubts on the interest of this proposal.

The negation criterion has been largely invoked in the analysis of causative constructions and serial verb constructions, but very rarely (if ever) in discussions about the mono- or biclausal nature of focus constructions. According to this criterion, focus constructions that have only one possibility of forming negation can be analyzed as monoclausal, whereas a monoclausal analysis is ruled out for the focus constructions that have two possibilities of forming negation, one of them similar to the negation of the corresponding plain assertive clause, the other similar to the negation of an identificational clause, as in (23).

(23) French

(23a) *C'est à Jean que j'ai parlé.*

it\_is to John that I\_have spoken

'It is to John that I spoke.'

(23b) *C'est à Jean que je n'ai pas parlé.*

it\_is to John that I NEG-have NEG spoken

'It is to John that I didn't speak.'

(23c) *Ce n'est pas à Jean que j'ai parlé.*

it NEG-is NEG to John that I\_have spoken

'It's not to John that I spoke.'

(23d) *Ce n'est pas à Jean que je n'ai pas parlé.*  
 it NEG-is NEG to John that I NEG-have NEG spoken  
 'It's not to John that I didn't speak.'

(24) illustrates the case of a focus construction that does not have this double possibility of forming negation, and consequently must be analyzed as monoclausal. In Mandinka, there is no way to express negative focalization within the frame of the focus construction illustrated in (24a), either by adding a negative marker, or by replacing *lè* by a morpheme expressing negation of identification. The only possibility is to use of a plain cleft construction, as in (24c), whose (24b) is the positive counterpart.

(24) Mandinka (pers.doc.)

(24a) *Áadámá lè yè Fàatú fútúu.*  
 Adama FOC CPL.TR Fatou marry  
 'ADAMA married Fatou.'

(24b) *Miŋ yè Fàatú fútúu, wõo mù Áadámá lè tí.*  
 REL CPL.TR Fatou marry DEM COP Adama FOC POSTP<sup>5</sup>  
 'The one who married Fatou, it's Adama.'

(24c) *Miŋ yè Fàatú fútúu, wõo tè Áadámá tì.*  
 REL CPL.TR Fatou marry DEM COP.NEG Adama POSTP  
 'The one who married Fatou, it's not Adama.'

Example (25) illustrates the two possible ways of forming negation in the focus construction of Wolof, either by suffixing the negative marker *-ul* to the verb, or by combining the clefted constituent with *du*, also found in the function of negative predicator in equative predication.

(25) Wolof (Diouf 2003 and Anna Marie Diagne pers.com.)

(25a) *Bind-uma Omar.*  
 write-NEG.1SG Omar  
 'I didn't write to Omar.'

(25b) *Omar laa bind-ul.*  
 Omar FOC.1SG write-NEG  
 'It's to Omar that I didn't write.'

(25c) *Du Omar laa bind.*  
 it\_is\_not Omar FOC.1SG write  
 'It's not to Omar that I wrote.'

<sup>5</sup> In the equative predication construction of Mandinka, the noun phrase expressing identification is marked by a postposition also found in other contexts as the equivalent of English 'as'.

Insofar as information about the interaction of focalization and negation can be found in the available descriptions (which unfortunately is not often the case), it seems that morphologically marked focus constructions with a double possibility of forming negation are more common than those devoid of this possibility, even among focus constructions that otherwise show evidence of monoclausality, as for example in French, where the two possibilities of forming negation in cleft constructions suggests a biclausal analysis, whereas the agreement of the verb with the clefted constituent supports a monoclausal analysis.

Similarly, the focus construction of Zarma illustrated in (3c) above (repeated here as (26a)) shows no other evidence of biclausality than the possibility of negating the focused constituent by means of the negative marker *màncí*, used specifically in identificational clauses.

(26) Zarma (Oumarou Yaro 1993: 108-111)

(26a) *Ábdù nòo Múusà díi.*  
Abdou it\_is Moussa see  
'Moussa saw ABDOU.'

(26b) *Màncí Ábdù nòo Múusà díi.*  
NEG Abdou it\_is Moussa see  
'It's not Abdou that Moussa saw.'

Consequently, Haspelmath's proposal of considering the negation criterion as the only one really decisive in the evaluation of the mono- or biclausal nature of a construction is somewhat problematic for the analysis of cleft constructions. The observation of the evolution of clefts rather suggests that bi- vs. monoclausality is a gradient rather than categorical distinction, and that, in the trend towards monoclausality that unquestionably characterizes the evolution cleft constructions, the behavior of clefts with respect to negation is particularly resistant to change.

Moreover, the double possibility of forming negation can be observed even in focus constructions that involve no specific morphological material at all, and consequently do not show evidence of resulting from the evolution of a cleft construction, as in Hungarian – example (27).

(27) Hungarian (É. Kiss 2002: 4)

*János nem MARI-T nem hívta fel.*  
John NEG Mary-ACC NEG called up  
'As for John, it was not Mary that he did not call up.'

## 8. Conclusion

In this paper, after clarifying the terminology (section 2) and briefly commenting the specificity of clefts in comparison to other subordinating constructions (section 3), I have reviewed the types of processes that typically occur in the history of clefts (section 4), and

commented the emergence of focus markers as a possible outcome of the evolution of cleft constructions (section 5). Section 6 briefly addressed the question of semantic changes that may result in a mismatch between the formal articulation ‘clefted constituent / rest of the construction’ and the semantic articulation ‘presupposition / assertion’. Finally, in section 7, I have tried to draw the attention to the issue of the negation-focalization interplay, and to the possible use of negation as a criterion of mono- vs. biclausality in the analysis of clefts. On this last point, it turns out that, although there is a consensus on the trend towards monoclausality in the evolution of cleft constructions, the focus constructions that result from the grammaticalization of a cleft construction tend to maintain a double possibility of forming negation, even when other criteria clearly support a monoclausal analysis.

## References

- Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry* 1. 149-168.
- Ball, C.N. 1991. *The historical development of the it-cleft*. University of Pennsylvania dissertation.
- Bohnemeyer, Jürgen, Nicholas J. Enfield, James Essegbey, Iraide Ibarretxe-Antuñano, Sotaro Kita, Friederike Lüpke & Felix K. Ameka. 2007. Principles of event segmentation in language: the case of motion events. *Language* 83(3). 495–532.
- Clech-Darbon, Anne, Annie Rialland & Georges Rebuschi. 1999. Are there cleft sentences in French? In Laurie Tuller & Georges Rebuschi (eds.), *The grammar of focus*. John Benjamins. 83-118.
- Comrie, Bernard. 1995. Serial verbs in Haruai (Papua New Guinea) and their theoretical implications. *Langues et langage: Problèmes et raisonnement en linguistique: Mélanges offerts à Antoine Culioli*, in Janine Bouscaren, Jean-Jacques Franckel & Stéphane Robert (eds.). Paris: Presses Universitaires de France. 25–37.
- Creissels, Denis & Pierre Sambou. 2013. *Le mandinka: Phonologie, grammaire, textes*. Paris: Karthala.
- Diouf, Jean-Léopold. 2003. *Dictionnaire wolof-français et français-wolof*. Paris: Karthala.
- Dufter, Andreas. 2008. On explaining the rise of *c’est*-clefts in French. In Ulrich Detges & Richard Waltereit (eds.), *The paradox of grammatical change: Perspectives from Romance*. John Benjamins. 31-56.
- Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
- Güldemann, Tom. 2010. The relation between focus and theticity in the Tuu family. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa*. John Benjamins. 69-93.
- Gundel, J. K. 1977. Where do cleft sentences come from. *Language* 53. 543-549.
- Haegemann, Liliane, André Meinunger & Aleksandra Vercauteren. 2014. The architecture of *it*-clefts. *Journal of Linguistics* 50(2). 269-296.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge University Press.
- Haspelmath, Martin. 2016. The serial verb construction: comparative concepts and cross-linguistic generalizations. *Language and Linguistics* 17(3). 291-319.

- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg : Helmut Buske.
- Kato, Mary A. & Carlos Miotto. 2016. Pseudo-clefts and semi-clefts. In Mary A. Kato & Francisco Ordóñez (eds.), *The Morphosyntax of Portuguese and Spanish in Latin America*. New York: Oxford University Press. 286-306.
- Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar*. Yale University Press.
- Oumarou Yaro, Bourahima. 1993. *Eléments de description du zarma*. PhD dissertation. University of Grenoble.
- Patten, Amanda L. 2012. *The English it-cleft : A constructional approach and a diachronic investigation*. Mouton de Gruyter.
- Robert, Stéphane. 2010. Focus in Atlantic languages. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa*. John Benjamins. 233-260.
- Sasse, Hans-Jürgen. 1985. Thethetic/categorical distinction revisited. *Linguistics* 25(3). 511-580.
- Van der Wal, Jenneke & Jacky Maniacky. 2015. How 'person' got into focus. *Linguistics* 53(1).1-52.
- Zentz, Jason. 2016. The biclausal status of Shona clefts. 90th Annual Meeting of the Linguistic Society of America.
- Zubizarreta, María Luisa. 2014. On the grammaticalization of the assertion structure: A view from Spanish. In Andreas Dufter & Álvaro S. Octavio de Toledo y Huerta, *Left sentence periphery in Spanish: diachronic, variationist and comparative perspectives*. John Benjamins. 253-282.