

THE VIETNAMESE POLYFUNCTIONAL MARKER *mà* AS A GENERALIZED LINKER: A MULTILEVEL APPROACH

Danh Thành DO-HURINVILLE
Université de Franche-Comté, Besançon & ELLIADD EA4661
<danh_thanh.do-hurinville@univ-fcomte.fr>

Huy Linh DAO
INALCO & CRLAO UMR 8563 (CNRS)
<huy-linh.dao@inalco.fr>

Abstract

The main aim of this paper is to illustrate the notions of polyfunctionality and transcategoriality as described in Robert (2003, 2004) through the case study of the marker *mà* in modern Vietnamese. We lay out a multilevel analysis to account for its flexible categorial status. It will be proposed that *mà* is best treated as a Generalized Linker whose basic function consists of connecting two constituents of different types. We show that the transcategorial behavior of this morpheme results from its ability to operate on different levels: phrase level (relativizer, verbal conjunction), clause level (correlative conjunction), and sentence level (attitudinal sentence-final particle, discourse connector). In our account, the uses of *mà* at the sentence level and the weakening of its core function are the byproduct of some form of ellipsis targeting one of the two constituents connected by *mà*.

Keywords: *Mà*, generalized linker, Vietnamese Grammar, transcategoriality, relativizer
ISO 639-3 codes: vic

1 Introduction¹

Robert et al. (2003), dealing with research on African isolating languages, argue that a variable portion of morphemes in those languages can be used synchronically across different syntactic categories. Linguistic units exhibiting such a behavior illustrate what one may call “synchronic grammaticalization” or “polygrammaticalization” (Craig 1991) and are referred to by these authors as “transcategorial morphemes”. As they emphasize, the categorial flexibility of lexical items is far from being restricted to African languages and actually can be observed cross-linguistically. Robert (2003, 2004) shows at great length that transcategoriality is a general property of linguistic systems which is exploited by languages in a variety of ways. For her and her colleagues, “transcategoriality constitutes, in all likelihood, a means of optimization of linguistic systems, allowing a minimum of forms to have a maximum of functions” (Robert 2003:18).²

In this respect, Robert et al.’s viewpoint is similar to that of Anward (2000:38).

As we have seen, languages tend to make optimal use of their lexical resources. Instead of coining distinct items for every combination of concept and function, languages tend to recycle items in several functions. But new functions of old items must be identifiable. This means that each language must strike a balance between

¹ We would like to thank the anonymous reviewer for his/her constructive comments and suggestions that helped us to improve the quality of our manuscript. As a normal disclaimer, errors are of course of the authors.

² The translation is ours. The original text in French is “la transcategorialité constitue vraisemblablement un moyen d’optimisation des systèmes linguistiques, permettant à un minimum de formes d’avoir un maximum de fonctions”.

flexibility (recycling) and contrast (identification), and such balances tend to block complete recycling of all items [...], language users strive to maximise meaning and minimise effort.

(Anward 2000:38)

Robert et al.'s stance ("a minimum of forms with a maximum of functions") and Anward's ("maximise meaning and minimise effort") correspond to Zipf (1949)'s "Principle of Least Effort" and Martinet (1955)'s "Principle of Linguistic Economy", which, in the human activity, recognize two paradoxical strengths which satisfy "communicative needs" and limit "the effort".

The polyfunctionality of Vietnamese morphemes such as *mà* (relativizer, contrastive conjunction, sentence-final particle, etc.), *thì* (topicalizer, sequential conjunction, etc.) and *là*³ (copula, speech verb relativizer, sentence-final particle, etc.) illustrate the positions of Robert et al., Anward, Zipf and Martinet. In this paper, we will focus on the Vietnamese polyfunctional marker *mà* and proceed as follows. We first present previous approaches to this marker (§2). We then put forward our own account, whereby *mà* is treated as a Generalized Linker (§3) which is used to connect two constituents of different types. Crucially, our analysis is based on the idea that the various functions fulfilled by this morpheme, as well as its various categories, are determined (for the most part) by the contexts it occurs in. As such, the categorial status of *mà* depends upon the level on which it operates. It can thus be analyzed as a relativizer or a verbal conjunction at the phrase level (§4); as a correlative conjunction at the clause level (§5); and as a final particle or a discourse connector at the sentence level (§6).

2. Previous approaches to *Mà*

There are two main approaches to *mà*: one is syntactic and the other is pragmatic. They can be summarized as follows.

2.1 Syntactic approach (Thompson 1965)

According to the syntactic approach taken in Thompson's (1965) grammar, *mà* is analyzed as a coordinating particle (example 1), a descriptive complement particle (example 2), or a final particle (example 3).

- (1) *Tôi muốn được gặp ông ấy, mà ông đi rồi* (Thompson)
 1S want get meet 3S **mà** 3S leave CRS
 'I wanted to get to meet him, but he's left already.'

- (2) *Tôi đã tìm thấy quyển sách mà anh nói hôm nọ* (Thompson)
 1S PFV find see CL book **mà** 2S talk day other
 'I found the book you were talking [about] the other day.'

- (3) *Ông ấy đã cưới vợ lâu rồi mà!* (Thompson)
 3S PFV married long time CRS **mà**
 'Why, he's been married for a long time [contrary to what you suggest]!'

In his original work, Thompson did not provide a thorough treatment of the second use of *mà* (ex. 2). It is not clear whether what he called "a descriptive complement particle" should be analyzed as a relative pronoun, a complementizer or merely a subordinating conjunction. Throughout the remainder of this paper, we refer to *mà* in this use as a relativizer, term used by Amara & Natchanan (2008) in their typological study of relative clauses in mainland Southeast Asian languages. The sequence introduced by *mà* in (2), namely *mà anh nói hôm nọ* '(that) you were talking about the other day', is therefore considered a relative clause. Since we are mainly concerned with illustrating the polyfunctional behavior of *mà*, we do not attempt to go further into the complex question of the structure of this type of clauses.

³ *Thì* and *Là* derive from lexemes: *thì* is a Sino-Vietnamese term 時 meaning "time" (see Alves 2007:221), *là* is a copular verb. As markers, *thì* functions as a temporal anaphoric conjunction or a topic marker, whereas *là* functions as a consecutive conjunction, a focus marker or a complementizer.

2.2 Pragmatic approach (Cao 2004)

In reaction to the syntactic approach, according to which Vietnamese is a subject-predicate language, and in accordance with the classification of languages by Li Ch. and Thompson (1976), Cao (2004) states that Vietnamese is a Topic-Prominent language. He adopts a pragmatic approach and argues that *mà* is a pragmatic marker used to separate the sub-topic from the sub-rheme of the topic, as in (4) below. According to Cao, in (4), the topicalizer *thì* separates the topic *bây giờ mà đi bộ* from the rheme *không kịp*. Within this topic, the marker *mà* is used to separate the sub-topic *bây giờ* ‘now’ from the sub-rheme *đi bộ* ‘walk’.

- (4) *Bây giờ mà đi bộ thì không kịp.*
 now mà walk TOP NEG in time
 ‘If we walk now, we won’t be in time.’

In Cao’s perspective, sentences containing a restrictive relative clause such as (5) can be described in the same fashion. In (5), the complex NP *Cuốn sách mà tôi đang đọc* ‘the book that I am reading’ constitutes the topic of the sentence, while the rheme is expressed by the predicate *rất hay* ‘is very interesting’. Here again, *mà* serves to divide the topic into a sub-topic (*Cuốn sách* ‘the book’) and a sub-rheme (*tôi đang đọc* ‘I am reading (it)’).

- (5) *Cuốn sách mà tôi đang đọc // rất hay.*
 CL book mà 1S prog read // very interesting
 ‘The book that I am reading is very interesting.’

Cao’s pragmatic approach can be said to be close in spirit and similar to the notion of “communicative dynamism” (CD), one of the basic concepts of the functional sentence perspective (FSP) theory developed by Firbas (1971:135-136). This concept is based on the idea that linguistic communication is not a static, but a dynamic phenomenon.

By CD I understand a property of communication, displayed in the course of the development of the information to be conveyed and consisting in advancing this development. By the degree or amount of CD carried by a linguistic element, I understand the relative extent to which the element contributes to the development of the communication, to which, as it were, it ‘pushes the communication forward’.

According to this view, the elements bringing the lowest degree of CD constitute the topic, those bringing the highest degree of CD are the rheme. Thereby, the sentence *He was cross* could be interpreted in regard to the degrees of CD as follows. The lowest degree of CD is carried by *He*, the highest by *cross*, the degree carried by *was* ranking between them (Firbas:136).

In general, the sequence preceding *mà*, viz. the sub-topic, must be less rhematic than the one following it, viz. the sub-rheme. This is indeed the case in (5): *tôi đang đọc* ‘I am reading (it)’ is “heavier” than *cuốn sách* ‘the book’ in terms of communicative dynamism.

3. *Mà* as a generalized linker

In what follows, we attempt to show that *mà* is fundamentally a generalized linker⁴, in that it exhibits such a behavior in all the environments it appears in. That is, *mà* remains a linker in all of its uses, even when it is analyzed as a final particle. As such, *mà* illustrates perfectly the notions of polyfunctionality and transcategoriality described in Robert et al. (2003). More precisely, *mà* is a kind of conjunction which serves to connect two constituents A and B, which, as we will see, are not always obligatorily overtly realized. The table 1 below summarizes the five possible cases of the constituents A and B.

⁴ This view is neutral with respect to the syntax vs. pragmatics discussion.

Table 1: Types of constituents and different levels of linkage by *mà*

Constituent A		Constituent B	Level
Case 1	Noun phrase	Clause or verb phrase	Phrase
Case 2	Verb phrase	Verb phrase	Phrase
Case 3	Subordinate clause	Main clause	Clause
Case 4	Clause	∅	Sentence
Case 5	∅	Clause	Sentence

The five cases above show that *mà* can be used at the three following syntactic levels: phrase, clause and sentence. At the phrase level, *mà* functions as a relativizer, which is a special type of conjunction (case 1), or as a verbal conjunction connecting two VPs (case 2). At the clause level, *mà* functions as a correlative conjunction (case 3). At the sentence level, *mà* is placed at the end of the sentence and is called a final particle (case 4). It is used as a modal particle expressing the speaker’s attitude. However, at the same (sentence-) level, when it appears sentence-initially, it can be characterized as a discourse connector whose function consists in signaling relations between prior and present discourse (case 5). In other words, it signals “the kind of relations a speaker perceives between different parts of the discourse” (Lenk 1997:2). The table 2 below summarizes the polyfunctionality (or the transcategoriality) of *mà*.

Table 2: Polycategoriality of *mà*

Gram	Target	
<i>mà</i>	Syntactic levels	Categories
	Level 1	Relativizer or verbal conjunction
	Level 2	Correlative conjunction
	Level 3	Final particle or discourse connector

4. *Mà* at phrase level

Mà is used at the phrase level when it functions as a relativizer that connects a NP to a clause or to a VP (cf. case 1), or as a verbal conjunction which connects two VPs (cf. case 2).

4.1 *Mà* as a relativizer

Based on a close examination of several examples containing *mà*, we note that the constituent A, i.e. the noun phrase (NP) preceding it, can be indefinite, definite, or generic.

4.1.1 Indefinite NP

The constituent A can be interpreted as assuming various grammatical functions which correspond to different syntactic positions within the constituent B, such as subject, direct object, indirect object, or possessor, a.o. In this case, A is a NP whereas B must be a clause or a VP. According to Keenan and Comrie (1977), there are six types of “Hierarchically arranged relativized noun phrases” as follows: Subject > Direct Object > Indirect Object > Oblique > Possessor > Object of Comparison. Prasithratsint and Yaowapat (2008:9) state that this “hierarchy implies that some nouns are more accessible or easier to relativize than other nouns. The sign > in the hierarchy means more accessible to relativization. So subjects are more accessible to relativization than direct objects, direct objects are more accessible to relativization than indirect objects, and so on”.

The relativization of subject, direct object, indirect object and possessor will be illustrated in (6).

a. Constituent A = subject

(6) *Người đàn ông mà đang đọc sách là thầy tôi.*
 person man **mà** PROG read book COP professor 1S
 ‘The man who is reading a book is my professor.’

(6a) **Người đàn ông ----- là thầy tôi.*

(6b) *Người đàn ông Ø đang đọc sách là thầy tôi.*

The NP (= A) *người đàn ông* ‘man’ is made definite by the use of the VP (= B) *đang đọc sách* ‘is reading’, introduced by the relativizer *mà*. As shown in (6a), the restrictive relative clause cannot be deleted while *mà* is omissible, as in (6b). According to Cao’s pragmatic approach, *người đàn ông mà đang đọc sách* ‘The man who is reading a book’ is the topic whereas *là thầy tôi* ‘is my professor’ is the rheme. Within the topic, *mà* serves to separate the sub-topic *người đàn ông* from the sub-rheme *đang đọc sách*.

b. Constituent A = direct object

(7) *Cuốn sách mà tôi đang đọc rất hay.*
 CL book mà 1S PROG read very interesting
 ‘The book (that) I am reading is very interesting.’

(7a) *Cuốn sách Ø tôi đang đọc rất hay.*

In (7), the relative clause *mà tôi đang đọc* ‘that I am reading’ restricts the reference of the NP *cuốn sách* ‘the book’, and therefore makes it definite. The relativizer *mà* can be omitted as in (7a). From the pragmatic viewpoint, *cuốn sách mà tôi đang đọc* ‘the book that I am reading’ is the topic of the whole sentence and *rất hay* ‘is very interesting’ is the rheme. Within the topic, *mà* separates the sub-topic *cuốn sách* ‘book’ from the sub-rheme *tôi đang đọc* ‘I am reading’.

c. Constituent A = indirect object

(8) *Đứa bé mà tôi cho tiền hồi nãy là cháu tôi.*
 CL child mà 1S give money moment ago COP nephew 1S
 ‘The child to whom I gave some money a moment ago is my nephew.’

(8a) *Đứa bé Ø tôi cho tiền hồi nãy là cháu tôi.*

In (8), the NP *đứa bé* ‘child’ is made definite by the relative clause *mà tôi cho tiền hồi nãy* ‘that I gave some money a moment ago’. Here again, *mà* is optional, as evidenced in (8a).

d. Constituent A = possessor

(9) *Cô gái mà chiếc xe đạp bị hư là em gái tôi.*
 young woman mà CL bicycle broken down COP young sister 1S
 ‘The young woman whose bicycle is broken down is my younger sister.’

(9a) **Cô gái Ø chiếc xe đạp bị hư là em gái tôi.*
 young woman CL bicycle broken down COP young sister 1S
 ‘The young woman who has a broken down bicycle is my younger sister.’

(9b) *Cô gái có chiếc xe đạp bị hư là em gái tôi.*
 young woman has CL bicycle broken down COP young sister 1S
 ‘The young woman having a broken down bicycle is my younger sister.’

(9c) *Cô gái mà có chiếc xe đạp bị hư là em gái tôi.*
 young woman mà has CL bicycle broken down COP young sister 1S
 ‘The young woman who has a broken down bicycle is my younger sister.’

In (9), the NP *cô gái* ‘young woman’, which is interpreted as the possessor of *chiếc xe đạp* ‘the bicycle’, is identified and rendered definite by the relative clause *mà chiếc xe đạp bị hư* ‘whose bicycle is broken down’. Unlike examples (6), (7) and (8), the use of *mà* is obligatory in (9). As a result, (9a) is not felicitous if *mà* is missing. Nonetheless, it is possible to use the verb *có* ‘have’, as in (9b), to make explicit the relationship of possession. The relativizer *mà* and the verb *có* can cooccur as in (9c).

4.1.2 Definite NP

The constituent A can also be an indexical expression, which is inherently definite. In that case, it can function as subjects (10) or as temporal adverbials (11). Using *mà* can generate a hypothetical value as shown in (10) and (11). This observation is consistent with Hagege’s viewpoint (2001:61), who specifies that “the relative clause is a determinant which contains a predicate, and can even express a hypothetical value”⁵. Note, however, that in this case and the one discussed in § 4.1.3, *mà* seems to behave less like a relativizer than a pure linker or a non-assertive marker (see Dao & Do-Hurinville 2013 for a discussion).

a. Constituent A = subject

(10) *Anh mà không giúp nó thì việc ấy hỏng.*
 2S mà NEG help 3S TOP work DEM incompleted
 ‘If you don’t help him, that work won’t be completed.’

(10a) *Anh Ø không giúp nó thì việc ấy hỏng.*

b. Constituent A = temporal adverbial

(11) *Ngày mai mà anh đi thì trễ rồi.*
 tomorrow mà 2S leave TOP late CRS
 ‘If you leave tomorrow, it will already be too late.’

(11a) *Ngày mai Ø anh đi thì trễ rồi.*

In (10) and (11), since the NPs *anh* ‘you’ and *ngày mai* ‘tomorrow’ are inherently definite, the relative clauses are no longer necessary and become non-restrictive. Consequently, they can give rise to a hypothetical reading. The deletion of *mà* is possible as in (10a) and (11a). However, when present, it allows to emphasize the hypothetical meaning. Consider the following examples (12) to (12e), in which the NP *trời* (sky) is definite.

(12) *Trời Ø mưa thì ta ở nhà.*
 sky rain TOP 1P stay home
 ‘If it rains, we stay home.’

(12a) *Trời mà mưa thì ta ở nhà.*
 sky mà rain TOP 1P stay home
 ‘If it rains, we stay home.’

(12b) *Nếu trời ---- mưa thì ta ở nhà.*
 if sky rain TOP 1P stay home
 ‘If it rains, we stay home.’

⁵ Translation ours, the original text being “la relative est un déterminant qui contient un prédicat, et peut même valoir une hypothétique”.

(12c) *Nếu trời mà mưa thì ta ở nhà.*
 if sky mà rain TOP 1p stay home
 ‘If it rains, we stay home.’

(12d) *Nếu mà trời mưa thì ta ở nhà.*
 if mà sky rain TOP 1P stay home
 ‘If it rains, we stay home.’

(12e) *Nếu mà trời mà mưa thì ta ở nhà.*

The examples (12) to (12e) are all truth-conditionally equivalent and they all generate hypothetical meaning. In each of these examples, the marker *thì* connects two clauses: the first clause expresses the condition, the second one indicates the consequence. The fact that (12) is well-formed and is able to convey a hypothetical statement without *mà* seems to suggest that the definiteness of the initial NP plays a key role in the emergence of modal meaning here. In (12a), *mà* separates the subject *trời* ‘sky’ from the predicate *mưa* ‘rains’. Example (12b) illustrates the most frequently referred to conditional construction in Vietnamese, whereby the first clause is introduced by the subordinator *nếu* ‘if’. Interestingly, as we can see in (12c), *mà* and *nếu* can coocur in the same sentence, with the NP subject *trời* ‘sky’ being sandwiched between them. *Mà* can also be used simultaneously with *nếu* to form a complex subordinator *nếu mà* as in (12d). Evidence in support of this latter hypothesis comes from the possibility for *mà* to appear twice, the second occurrence being placed in its usual position, as in (12e).

That being said, examples (12) to (12d) are not equivalent from the pragmatic viewpoint. The juxtaposition of the two clauses connected by *thì* in (12) primarily conveys a sense of condition (first clause) and of consequence (second clause). In the presence of *mà* in (13a), the emphasis seems to be put chiefly on the predicate *mưa* ‘rain’, which is highlighted as the condition. The hypothetical meaning is rendered explicit by *nếu* (if) in (12b) rather than by *mà* in (12a). The difference between (12c) and (12d) is that *mà* has scope over the predicate *mưa* ‘rain’ in (12c), while this marker scopes over the whole clause *trời mưa* ‘it rains’ in (12d).

4.1.3 Generic or kind-referring NP

This type of NPs is akin to the definite NPs discussed above to the extent that they seem to display the same effects with regard to the emergence of hypothetical meaning. Here again, when preceded by kind-denoting NPs, the sequences introduced by *mà* can no longer act as restrictive modifiers. In the two following examples, the constituent A is a generic NP that functions as a subject.

(13) *Chó mà không biết sủa thì không phải là chó.*
 dog mà NEG know bark TOP NEG must COP dog
 ‘A dog that cannot bark is not a dog.’
 ‘If a dog cannot bark, it cannot be a dog.’

(14) *Mèo mà không biết bắt chuột thì không phải là mèo.*
 cat mà NEG know catch mouse TOP NEG must COP cat
 ‘A cat that cannot catch mice is not a cat.’
 ‘If a cat cannot catch mice, it cannot be a cat.’

Examples (13) and (14) are generic statements. Because of the absence of the animate classifier *con*, whose function is to return a count noun when combined with a bare noun, *chó* ‘dog’ and *mèo* ‘cat’, being mass nouns, refer to the species *chó* and *mèo*. The relative clauses in (13) and (14) generate a hypothetical value, as in (10), (11) and (12a). The following table summarizes the function of *mà* in (6) to (14).

Table 3: *Mà* as a relativizer

Constituent A	Relativizer	Constituent B
NP (indefinite, definite, generic)	<i>Mà</i>	Clause or VP

4.2 *Mà* as a verbal conjunction

If *mà* can connect a NP to a clause (or a VP) as in (6) to (14), it can also connect two VPs. When *mà* links two VPs, it functions as a verbal conjunction, and is mandatorily realized. Let us examine the following examples:

- (15) *Nhà hàng này ngon mà rẻ lắm.*
 restaurant DEM delicious **mà** cheap very
 ‘In this restaurant, the food is delicious but very cheap.’

- (15a) *Nhà hàng này ngon và rẻ lắm.*
 restaurant DEM delicious and cheap very
 ‘In this restaurant, the food is delicious and very cheap.’

In (15), *mà* links two stative predicates *ngon* ‘be delicious’ and *rẻ* ‘be cheap’ which occur concomitantly. The speaker states that the restaurant possesses two qualities, viz. “be delicious” and “be cheap”. In this context, *mà* is interchangeable with the conjunction *và* ‘and’ as in (15a). However, the use of *mà*, which focuses on the second VP, indicates that *rẻ* ‘be cheap’ constitutes an additional and more important quality of the restaurant (with respect to *ngon* ‘be delicious’). From the pragmatic viewpoint, *mà*, unlike *và*, conveys a conventional implicature, namely that there is a contrast between being delicious and being cheap, since delicious things are normally expected to be expensive. By choosing to put forth, in the first place, the fact that the restaurant offers delicious food, the speaker probably tried to encourage the addressee to go there. Nonetheless, knowing about that may lead one to infer that the price would be high and it might not be a good idea to go there. The effect of the second argument, the unexpected cheapness of the price, is to cancel this inference and to allow the speaker to give more persuasive power to his/her suggestion (not only “is it delicious” but it “is also very cheap”).

- (16) *Nhà hàng này ngon mà đắt lắm.*
 restaurant DEM delicious **mà** expensive very
 ‘In this restaurant, the food is delicious but very expensive.’

- (16a) *Nhà hàng này ngon ??và đắt lắm.*
 restaurant DEM delicious and expensive very
 ‘In this restaurant, the food is delicious but very expensive.’

Interestingly enough, the stative verb *rẻ* ‘be cheap’ can be replaced with the antonymous stative verb *đắt* ‘be expensive’ as in (16). Here, the argumentative orientation of the utterance is no longer the same, even though the concessive value of *mà* remains unchanged. Contrary to (15), the speaker in (16) might intend to convince his/her interlocutor not to go to the restaurant. By saying what is expected to be inferred from the first property (being delicious normally means being expensive), the speaker validates the inference made by the hearer. While there is still a contrast between the properties denoted by the two VPs, this contrast operates on different kinds of inferences:

- In (15), it is established on the basis of the inferences (delicious → inference1: expensive → inference2: negative >< positive: inference3 ← cheap; intention: cancel the inference2 → overall outcome: positive);
- In (16), it operates directly on the basis of the stated qualities (delicious → inference1: positive >< negative: inference2 ← expensive; intention: cancel the inference1 → overall outcome: negative). The use of *và* ‘and’ in (16a) results in oddness.

- (17) *Chị ấy phải làm việc vất vả mà nuôi con.*
 3S must work hard mà feed children
 ‘She must work hard to feed her children.’

- (18) *(Anh) Lấy áo (tôi) mà mặc!*
 2S take shirt 1S mà put/wear
 ‘Take my shirt and put it on!’

In contrast to (15) and (16), *mà* in (17) and (18) connects two dynamic VPs: *làm việc vất vả* ‘work hard’ and *nuôi* ‘feed’ in (17); *lấy áo* ‘take shirt’ and *mặc* ‘put/wear’ in (18). When *mà* connects two stative VPs, it introduces an additional value represented by the second VP, as in (15) and (16). When it connects two dynamic VPs, the second VP is interpreted as the goal/purpose of the first VP, as in (17) and (18). The eventualities denoted by the two dynamic verbs can either occur simultaneously as in (17), or the second eventuality takes place logically after the first one as in (18). From the pragmatic viewpoint, *mà* can convey an imperative meaning as shown in (18).

- (19) *Cuốn sách này mà₁ anh thích thì giữ mà₂ đọc.*
 CL book DEM mà 2s like TOP keep mà read
 ‘If you like this book, keep it and read it.’

- (19a) *Cuốn sách này Ø anh thích thì giữ mà₂ đọc.*

Example (19) illustrates two different functions of *mà* within the same sentence. *Mà₁* acts as a relativizer and the whole sentence expresses a hypothetical meaning because the initial NP, namely *cuốn sách này* ‘this book’, is a definite NP. *Mà₂* is a verbal conjunction which connects two dynamic verbs *giữ* ‘keep’ and *đọc* ‘read’. Note that *mà₁* can be omitted whereas *mà₂* must be overtly realized, as shown in (19a).

- (20) I suggest *that₁ that₂* is the solution *that₃* we offer to the Yugo-Slavian police [...]
 (Agatha Christie, Murder on the Orient Express, Fontana, Collins, 1974, p. 191)

- (21) *Je suggère que₁ c’est la solution que₂ nous proposerons à la police yougoslave [...]*
 (translation ours)

A relativizer which cooccurs with a conjunction also exists in English. In the excerpt (20), *that₁* as complementizer can cooccur in the same sentence with *that₂* as demonstrative and with *that₃* as relativizer⁶. Notice that *mà₁* and *that₃* as relativizers are optional in Vietnamese and in English. In contrast, the use of the French relativizer *que₂* is obligatory as in (21).

5. *Mà* at clause level

In (15) to (19), *mà* functions as a verbal conjunction which connects two VPs of the same clause. This marker can function as a correlative⁷ conjunction (at the clause level) which connects a subordinate clause to a main clause (case 3) (see table 1) as illustrated in the following examples.

- (22) *Tại nó tới trễ mà tôi bị lỡ tàu.*
 due to 3S arrive late mà 1S suffer miss train
 ‘As he arrived late, I missed my train.’

⁶ *That₃* as relativizer is omitted in the original version.

⁷ Here, the term *correlative* refers to the fact that *mà* is regularly used in conjunction with another marker but is not adjacent to it.

- (23) *Vì trời mưa mà nó không đi học.*
 because sky rain mà 3S NEG go learn
 ‘As it was raining, he didn’t go to school.’
- (24) *Nhờ John giúp mà công việc của tôi trở nên dễ dàng hơn.*
 thanks to John help mà work POSS 1S become easymore
 ‘As John helped me, my work became easier.’

In (22), (23) and (24), subordinate clauses introduced by *tại, vì, nhờ* (a.o.) all express causality; main clauses introduced by *mà* indicate the consequence or the result of actions mentioned in subordinate clauses. One of the primary peculiarities of *mà* in this use is that, together with the subordinating conjunction, it forms a correlative system. As such, both elements are intimately connected, which requires that they be simultaneously and overtly realized. Omitting one of them thus leads to ungrammaticality. The following table summarizes the function of *mà* in (22), (23) and (24).

Table 4: *Mà* as a correlative conjunction

Subordinators	Constituent A	Correlative Conjunction	Constituent B
<i>Tại, vì, nhờ</i> (a.o)	Subordinate clause (Cause)	<i>Mà</i>	Main clause (Consequence)

6. *Mà* at sentence level

6.1 *Mà* as an attitudinal sentence-final particle⁸

Vietnamese and many other Asian languages depend to a great extent on the use of sentence-final particles (also called “attitudinal sentence-final particles”, cf. Li et al., 1982:21), which occur frequently in everyday conversations, dialogues in novels, television, or interviews in newspapers and magazines. When *mà* appears at the end of a sentence, it is often described as an “attitudinal” sentence-final particle (or a pragmatic particle) that indicates the speaker’s attitude towards the addressee. More precisely, the speaker can use *mà* in order to show her (usually negative) reaction towards the situation involving the addressee. For example, by uttering (25), the speaker means that she does not agree with her interlocutor, who seems to call into question the truth of what she has said. In this respect, *mà* clearly has a communicative function. Consider the following examples:

- (25) *Tôi nói thật mà!*
 1S speak true mà
 ‘But I speak the truth (believe me)!’
- (26) *John đã từng sống ở Việt Nam mà!*
 John PFV ever live LOC Vietnam mà
 ‘But John used to live in Vietnam!’
- (27) *Mẹ đã bảo rồi mà!*
 mum PFV tell CRS mà
 ‘But Mum has already told you so!’
- (28) *Trời ơi! Tôi tưởng anh đi rồi kia mà!*
 God oh 1S think 2S leave CRS FP mà
 ‘Oh God! (But) I thought (that) you had left!’

⁸ Interestingly, as the anonymous reviewer pointed out, neighboring Cantonese MA has a similar emphatic function (see Alves 2007:228).

- (29) *Tốt rồi mà!*
 okay CRS mà
 ‘But it’s OK (now)!’

Syntactically, *mà* occurs either at the end of sentences as in (25) to (28) or after a VP as in (29). In this latter case, the subject of the VP has arguably been dropped since its identity must be known to both the speaker and the hearer in the context of utterance of (29). The constituent A is the whole sentence, whereas the constituent B is missing. This use of *mà* can therefore be analyzed as stemming from some form of ellipsis, whereby *mà* initially acts as a conjunction coordinating two independent clauses, of which the second is elided. The absence of B means that A is directly connected to the speaker who is completely involved in his or her discourse.

It appears that *mà* can be readily deleted in all these examples without affecting their grammaticality or contextual autonomy, for the semantic contribution of *mà* is null and its conjunctive function is no longer justified in this context, the constituent B having been elided. However, from the pragmatic point of view, these statements sound incomplete and abrupt without *mà*, because this particle conveys the personal feelings and various kinds of attitudes of the speaker towards the hearer such as irritation, anger, explanation, persuasion, warning, reproach, surprise, etc. Put differently, *mà* performs a communicative function as in (25) to (29).

Without *mà*, (25) is only a neutral state of affairs. In the presence of this particle, the speaker tries to convince the hearer of his good faith and implicitly asks the latter to trust him or her. (26) is part of a dialogue and can be understood as a positive reaction of the current speaker in favor of the statement made by a previous speaker. This latter might have congratulated John on his Vietnamese skills. The current speaker then tried to express his or her agreement by adding an explanation that John speaks the language fluently because he lived in Vietnam before. In (27), the mother told her child not to play the fool, but the child did not listen to her. He fell down and began crying. In this situation, the mother uses *mà* to blame him for not following her advice. (28) contains three final markers: *rồi kia mà*. The marker *rồi* signals a “Currently Relevant State” (CRS) and always appears after the VP. The sentence-final particle *kia* which derives from the deictic *kia* ‘that’ reinforces the surprise of the speaker and must follow *rồi*. In (29), the speaker tries to convince the hearer that the issue has been taken care of and that the hearer does not have to worry at the speech moment T₀. In short, *mà* in (25) to (29) functions as a sentence-final particle and is used by the speaker to communicate different feelings to the hearer, such as persuasion in (25) and (30), explanation in (26), blame in (27), and surprise in (28). The following table summarizes the function of *mà* in (25) to (29).

Table 5: *Mà as an attitudinal particle*

Constituent A	Attitudinal particle	Constituent B
Sentence	<i>Mà</i>	∅

6.2 *Mà as a sentence-initial discourse connector*

The last use of *mà* we would like to present in this section shares with the previous one the property of operating at the sentence level and being observed preferably in spoken register. Nonetheless, it differs from the previous one in two respects. First, in this use, *mà* appears sentence-initially, which suggests that it is the constituent A, and not the constituent B, that is deleted. Second, unlike the final *mà*, the initial *mà* is better characterized as a discourse connector, in the sense that it serves to connect the present discourse (= B), i.e. the host-sentence, to the prior discourse (= A). The discursive link it establishes between them can be described as involving a topic shift, *viz.* it signals the point at which the speaker moves from one topic to another. More precisely, it indicates the intention of the speaker to close a topic in order to start a new one, as in (30), or to introduce a new idea, a new argument to those previously advanced, as shown in (31). The absence of *mà* in those examples does not affect their grammaticality. However, the discursive relationship between sentences must be inferred from the contexts of utterance.

- (30) Speaker 1:
Anh đã gặp bạn gái mới của Minh chưa? Xinh lắm!
 2S PFV meet girlfriend new POSS Nproper NEG pretty EXCL
 ‘Did you see Minh’s new girlfriend? She’s quite pretty!’

Speaker 2:
Rồi! Cũng được.
 Already also fine

Mà đây, anh quên không hỏi Minh vụ tài liệu rồi!
mà DEICT 1S forget NEG ask Nproper issue document CRS
 ‘I did! She’s fine. By the way/Speaking of which, I forgot to ask Minh about the document.’

- (31) *(Ở đây cũng dễ chịu lắm). Mà nó cũng thuận tiện nữa.*
 LOC here also pleasant EXCL **mà** 3S also convenient too
 ‘(Living here is so nice!) Besides, it’s also very convenient.’

The following table summarizes the function of *mà* in (30) and (31):

Table 6: Mà as a discourse connector

Constituent A	Discourse connector	Constituent B
(∅ = Prior discourse)	<i>Mà</i>	Current discourse

7. Conclusion

In this article, we have shown that the Vietnamese polyfunctional marker *mà* is best analyzed as a Generalized Linker whose function is basically to connect two constituents of various scales. In our account, its transcategorial behavior originates from its ability to operate on different levels. At the phrase level, it acts either as a relativizer or a verbal conjunction. At the clause level, together with another subordinator, it forms a correlative system. At the sentence level, it can appear either at the end of the sentence, in which case it functions as an attitudinal particle; or it can occur sentence-initially and fulfills the role of a discourse connector serving to signal a topic shift between prior and current discourses. When used as a linker, *mà* can be said to constitute a relative clause with the sequence following it. The definiteness of the NP preceding *mà* can have an impact on the interpretation of the relative clause, which is restrictive if the NP is indefinite, and non-restrictive if the NP is either definite or generic. In this latter scenario, the whole structure can convey hypothetical meaning. Employing *mà* as a verbal conjunction connecting two VPs allows the speaker to alter the argumentative orientation of his or her discourse, for *mà* can be associated with a contrastive conventional implicature besides its purposive value. It appears that the two constituents connected by *mà* are not always overtly realized. In our view, this possibility gives rise to the uses of *mà* as a final particle (ellipsis of the second constituent) or a discourse connector (ellipsis of the first one) at the sentence-level. One of the implications of this hypothesis is that the core function of *mà*, which is fundamentally that of a linker, is weakened when one or the other of the constituents is omitted⁹. Consequently, under such circumstances, *mà* becomes completely optional and can be deleted without resulting in ill-formedness of the host sentence which contains it¹⁰.

We would like to acknowledge, however, that at the current state of development of our work, we do not have historical data to support our view that all the uses described above historically derive from each other¹¹.

⁹ Nonetheless, the “linker” function is still present in those contexts, leaving the overt clause/sentence “hanging”, which leads to pragmatic interpretation.

¹⁰ This case is not the only one where *mà* can be omitted.

¹¹ As the anonymous reviewer rightly points out, the homophony of distinct words in the mental lexicon can potentially be related to other historical developments and chance similarities.

Rather, there appears to be one function as “linker”, which gets different interpretations in different semantic and syntactic contexts.

Abbreviations (Following the Leipzig glossing rules): 1s (first person singular); 1p (first person plural); 2s (second person singular); 3s (third person singular); CL (classifier); COP (copula); CRS (currently relevant state); DEICT (deictic); DEM (demonstrative); FP (final particle); EXCL (exclamative); LOC (locative marker); NEG (negative marker); PFV (Perfective marker); POSS (possessive marker); PROG (progressive marker); TOP (topicalizer).

References

- Alves, Mark J. 2007. Categories of grammatical Sino-Vietnamese vocabulary. *Mon-Khmer Studies* 37:217-229.
- Amara, Prasithratsint & Natchanan Yaowapat. 2008. A typology of relative clauses in mainland Southeast Asian languages. *Mon-Khmer Studies* 38:1-23.
- Anward, Jan. 2000. A dynamic model of part-of-speech differentiation. In Vogel & Comrie ed., *Approaches to the Typology of Word Classes*, 3-45. Berlin, New York: Mouton de Gruyter.
- Cao, Xuân Hạo. 2004. *Tiếng Việt, sơ thảo ngữ pháp chức năng*. Vietnam: Nhà Xuất Bản Giáo Dục.
- Craig (Grinevald), Colette 1991. Ways to go in Rama: a case study in polygrammaticalisation. In E. C. Traugott & B. Heine ed., *Approaches to grammaticalization*, 455-92. Amsterdam/ Philadelphia: John Benjamins.
- Dao, Huy Linh & Do-Hurinville, Danh Thành 2013. Les structures hypothétiques en vietnamien: *Mà* comme marqueur de non-assertion. *Travaux linguistiques du CerLiCo* 26.1:227-24.
- Do-Hurinville, Danh Thành 2013. La polyfonctionnalité et la transcatégorialité. Exemple de la conjonction *Mà* en vietnamien. *Langages* 190:101-118.
- Do-Hurinville, Danh Thành 2010. Les parties du discours en vietnamien: grammaticalisation et transcatégorialité. *Bulletin de la Société de Linguistique de Paris*, CIV/1:327-370.
- Firbas, Jan 1971. On the concept of communicative dynamism in the theory of functional sentence perspective, 135-144. *Sborník j/raci filologické fakulty brněnské university A* 19.
- Heine, Bernd & Kuteva Tania 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Keenan, Edward L. & Comrie Bernard 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63-99.
- Lenk, Uta 1997. Discourse markers. In J. Verschueren, J.-O. Östman, J. Blommaert, and C. Bulcaen ed., *Handbook of Pragmatics*. Amsterdam: John Benjamins.
- Li, Charles N., Thompson, Sandra A. & Thompson R. McMillan 1982. The Discourse motivation for the perfect Aspect: The Mandarin Particle *le*. In Paul J. Hopper ed., *Tense and Aspect*, 19-44. Amsterdam: John Benjamins.
- Li, Charles N. & Thompson, Sandra A. 1976. Subject and Topic: A New Typology of Languages. In Li Charles N. ed., *Subject and Topic*, 457-490. New York/San Francisco/London: Academic Press.
- Martinet, André 1955. *Economie des changements phonétiques. Traité de phonologie diachronique*. A. Francke Bern.
- Robert, Stéphane 2004. The challenge of polygrammaticalization for linguistic theory: fractal grammar and transcatégorial functioning. In Z. Frajzyngier, A. Hodges & D. S. Rood. ed, *Linguistic Diversity and Language Theories*, 119-142. Amsterdam/Philadelphia: John Benjamins.
- Robert, Stéphane 2003. Introduction de la grammaticalisation à la transcatégorialité. In Robert S. ed., *Perspectives synchroniques sur la grammaticalisation*, Afrique et Langage 5, 9-18. Louvain-Paris: Peeters.
- Thompson, Laurence C. 1967. *A Vietnamese Grammar*. Seattle & London: University of Washington Press.
- Vo, Thi Anh Ngoc. 2012. *Subordination relative en vietnamien: éléments de comparaison avec le français*. Doctoral Dissertation, Université de Toulouse, France.

Zipf, George K. 1949. *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.

Reviewed: Received 17 September 2018, revised text accepted 31 July 2019, published 15 September 2019

Editors: Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Nathan Hill, Dr Sigrid Lew