INCORPORATION
AND “FORMAL INCORPORATION”
IN ANALYTIC LANGUAGES:
MANDE LANGUAGES AND TYPOLOGY
OF INCORPORATION

Alexander Zheltov
St. Petersburg State University / Museum of Anthropology and Ethnography RAS
ajueltov@mail.ru

Abstract: Although the Mande languages are usually characterized as analytic, they
demonstrate a broader spectrum of typological features. For instance, Gban (South
Mande) was analyzed as having evident inflective elements. This paper concerns the
phenomenon of incorporation observed in some Mande languages (Mandinka, Tigemaxo, Soninke). It also attempts at attracting attention to the interpretation of
some facts which do not suit the definition of incorporation but demonstrate some
phenomena, to a certain extent, similar to incorporation (Gban, South Mande). The
author (very tentatively) uses the term “formal incorporation” for them and offers
a new variant of the typology of incorporation, including it into the broader set of
different phenomena. It is worth noting that the paper concentrates on the par-
ticular sort of incorporation, i.e., object + verb incorporation.

Key words: Mande, Gban, typology, incorporation, “formal incorporation”

1. Incorporation: general features

A common definition of incorporation is offered by Mithun (1984: 848): “The term incorporation is generally used to refer to a particular
type of compounding in which a V and N combine to form a new V.
The N bears a specific semantic relationship to its host V – as patient,
location, or instrument.” Mithun also points out that “all languages
which exhibit such morphological structures also have syntactic paraphrases. If we know that in Koryak, one can say *tiqoyanmatekn* ‘I-reindeer-slaughter’, then we can correctly predict the existence of a sentence like *tinmekin qoywge* ‘I-slaughter reindeer’” (Mithun 1984: 847). Though the term *incorporation* literally means the inclusion of something as part of a whole (in the sense of linguistics, for instance, object as part of a verb), this “inclusion” seems to be less important than the afore-mentioned parallelism and non-referentiality of an incorporated noun. In Mithun (2000: 916), it is noted that “incorporation is often found within polysynthetic languages, those with a high average number of morphemes per word. …but the two features are not necessarily linked... Many polysynthetic languages lack incorporation entirely, while some more analytic languages show it”. She gives some examples of the so-called “noun stripping”, which are considered as a sort of incorporation while the real inclusion of an object NP into the verb is not observed. In Tamil, such constructions keep all the features of incorporation (parallel constructions, object nouns do not appear with adjectives, number and case markers, and can affect the argument structure of the verb), but, at the same time, they do not form any phonological unit (the emphatic particle ‘indeed’ can occur between the stripped noun and verb) (Mithun 2000: 920–921).

So, incorporation usually combines transitive/intransitive (non-incorporated/incorporated) parallelism and phonological unity between object and verb in the latter case. It can also lack this phonological unity and still be treated as incorporation. This approach to incorporation appears to exclude from it the constructions demonstrating phonological unity between object and verb but having no parallel non-incorporated variant.

### 2. Incorporation in Mande

There are at least three Mande languages where incorporation was attested and analyzed: Mandinka (Creissels 2008; Creissels & Sambou 2013), Tigemaxo (Bozo) (Blecke 1996; 2011), Soninke (Creissels & Dramé
Analyzing the data concerning object + verb incorporation, which is the focus of this paper, we can see that in all three cases, we observe the “pure” incorporation in the sense of both the existence of parallel incorporated/non-incorporated forms and phonological object/verb compounding.

2.1. Mandinka

In Mandinka (Creissels 2008; Creissels & Sambou 2013: 303–310), we can see the incorporation of the subject (for the intransitive verb), the object (for the transitive verb), the nouns with oblique and comparative roles. It seems that for all the cases, there are parallel incorporated/non-incorporated variants, and the incorporated nouns form a closer than a purely syntactic unit with the verb: they lose definiteness, and insertion of something between them seems to be impossible. Concerning object + verb incorporation, it can be illustrated by the following examples where (1a) is the non-incorporated variant, while (1b) is the incorporated variant.

(1) Mandinka (Creissels 2008: 9)

a. À yé jío bòŋ
   3SG ACP.POSS water.DEF pour
   ‘He poured the water.’

b. À yé sålaatoo jíi-bøŋ
   3SG ACP.POSS salad.DEF water-pour
   ‘He sprinkled the salads.’ (lit. ‘He water-splashed the salads.’)

If these constructions are detransitivized, in the first case, it is ‘water’ that becomes a subject, while in the second case, the subject position is occupied by ‘salads’.  

---

1 In (Khachaturyan 2017), some features of N + V compounding are described in Mano. However, these features are more similar to Gban compounds discussed in the next sections than to the proper incorporation and are not treated as incorporation.
However, there is one “non-standard” trait: in the case of noun incorporation into the transitive verb, it does not become intransitive. Due to Mandinka verbs’ syntactic lability, this process looks similar to intransitive subject incorporation and is explained as a sort of possessor raising process.

2.2. Tigemaxo (Bozo)

As for Tigemaxo (Blecke 1996: 88–92; Blecke 2011), we can observe both incorporated (2b) / non-incorporated (2a) pairing and phonological indicators of compounding.

(2) Tigemaxo (Blecke 2011: 1)

a. \( n \text{ } ga \text{ } xulu \text{ } mana \)  
   \[ 1\text{SG} \text{ IPFV} \text{ pirogue} \text{ look_for[TR]} \]  
   ‘I look for a pirogue.’

b. \( n \text{ } ga \text{ } xulu-mɛnɛ \)  
   \[ 1\text{SG} \text{ IPFV} \text{ pirogue-look_for[INTR]} \]  
   ‘I “pirogue-look for”.’

In (2b), we can see that the incorporated verb is used in its intransitive form, while the non-incorporated alternation shows a transitive verb form (2a). It is different from the Mandinka examples but quite typical for “classical” incorporation. The compounding structure of N+V goes along with the loss of all sorts of determinatives for object and compact tonal scheme presented in (3).

(3) súà ‘matin’, xááí ‘montrer’ > súáxááí ‘saluer qn le matin’

It is also worth noting that if a verb has both a transitive and an intransitive form, then the incorporated variant has the intransitive verb form, i.e., the intransitivization of incorporated form is not optional, but not all verbs have two transitivity stems as it seems to be the case in (3).
2.3. Soninke

In Soninke (Creissels & Dramé 2018), incorporation can be observed in the possessive, object, and oblique constructions (similative, distributive constructions, constructions with temporal adjuncts). They can also form multiple incorporations of different combinations (temporal adjunct incorporation + object incorporation, etc.). For all of these types, there are parallel incorporated/non-incorporated forms. In the case of object incorporation, there is a special morphological marking for detransitivisation: “…marker -i can be used with all verbs ending with a, o, or u (example (4) – A.Zh.); with verbs ending with i or e, the antipassive marker -ndì may be used, but its use is optional. …In possessive incorporation and oblique incorporation, a linking (or epenthetic) -n- occurs between the two formatives of the compound verb (example (5) – A.Zh.)” (Creissels & Dramé 2018: 13).

(4) Soninke (Creissels & Dramé 2018: 13)
À wá yìràn-gáagé-né
3SG LOCCOP cloth-sell.DEPR-GER
‘(S)he sells clothes.’ or ‘She does cloth selling.’

(5) Soninke (Creissels & Dramé 2018: 13)
À bútítí-n-bí
3SG liver-EP-burn
‘(S)he got furious.’ (lit. ‘He/she liver-burnt.’)

So, in Soninke, incorporation is productive and has incorporated/non-incorporated pairings, in the case of object incorporation, it detransitivizes the verb and morphologically marks incorporated forms.

In general, we can conclude that all three languages represent paradigm cases of incorporation, as all of them show one of the essential features of incorporation, namely alternating incorporated / non-incorporated pairs. They also show different object/verb compounding types. Mandinka has an essential distinctive feature of non-detransitizing incorporated verbs. Soninke seems to be very productive in forming incorporated verbs, while Tigemaxo shows the highest formal
degree of N-V integration between object and verb (change in tonal contour). So, languages that in traditional typology are classified as analytic can indeed be incorporating.

3. “Formal incorporation” in Mande

Under the term “formal incorporation”\(^2\), I understand the phenomenon taking place in some Mande languages (especially South Mande): the constituents \([\text{Pronominal Subject} + \text{Aux (TAM)}] = \text{Pronominal Predicative Marker (PPM)} (\text{Vydrin 2010}) + \text{Object} + \text{Verb}\) have closer syntactic compactness / higher syntactic integration than other syntactic elements \((\text{Noun Subject, Indirect (Oblique) Object, Adjuncts})\).\(^3\) This fact is stated, for instance, by Nikitina (2011) who also mentions the absence of ditransitive constructions in Mande, non-typical for Niger-Congo word order \((\text{S-O-V-X})\) and the syntactic exclusion from VP of all constituents except the direct object. She explains this phenomenon by the analogy between NP and VP’s syntactic structures, which is typical for Mande and is confirmed by the regular formal coincidence between non-subject and possessive pronouns. So, VP with an object can be treated in the same way as NP with the inalienable possessor marker. Various formal means to join pronominal subject, object, and the verb are attested: tonal assimilation, contractive (portmanteau) forms \((\text{e.g., in Gban: PROSUBJ}+\text{TAM}, \text{PROSUBJ}+\text{PROOBJ}, \text{PROSUBJ}+\text{TAM} +\text{PROOBJ})\). Nominal subjects, oblique objects, and adjuncts are never involved in such processes. It makes the degree of syntactic independence of VP

\(^2\) In the first version of this paper, the term pseudo-incorporation was used instead of formal incorporation. One of the anonymous reviewers pointed out (with reference to Borik & Gehrke 2015) that in the recent literature, Mithun’s noun stripping is commonly referred to as pseudo-incorporation, and this term should be avoided concerning other phenomena. Taking into account the fact that phenomena under consideration are really different (or even opposite) to the notion of noun stripping, the term formal incorporation was introduced.

\(^3\) Similar phenomena are observed, for instance, in Gban (Zheltov 2008) and Mano (Khachaturyan 2017).
evidently lower than that of nominal subjects, oblique objects, adjuncts. If we consider the latter as separate lexemes, we should use a different term for the former ones, and it should not be the same as “separate lexemes”. In the case of a pronominal object (the subject can be either nominal or pronominal, since in languages with PPM, instead of personal pronouns, a PPM is used even if the subject noun is present), we can find the full chain of boundedness (PROSBJ-TAM-PROOBJ-V).

(6) Gban (personal data)

\[Mà \quad \text{bē} \]
\[1\text{SG.PRS.3SG} \quad \text{take.PRS}\]
‘I take him.’

In this short example, we can see many grammatical elements “glued” together: \(Mà = \tilde{t} (1\text{SG}) + \) low tone (grammatical tone for 1/2 person (locutors) pronouns in the Present tense) + \(a (3\text{SG.NSBJ}); \text{bē} = \) ‘take’ + low tone (the first syllable of the verb in the Present tense.

\footnote{In the case of pronominal arguments, S-TAM-O-V combination looks similar to the same combination in absolutely typologically different agglutinative Bantu languages, and the most remarkable difference is that Gban (despite being an analytic language) shows a much higher degree of fusion. Cf., for instance, example 6 with the same phrase in Swahili:

(i) Swahili (personal data)

\[Ni-na-m-chuku-a\]
\[1\text{SG.SBJ-PRS-3SG.OBJ-take-FV}\]
‘I take him.’

In both Swahili and Gban, nothing can be inserted between pronominal object and verbal stem, and the difference in glossing and writing can be explained, first, by tradition, and, second, by the fact that Swahili has S-V-O order for nominal arguments preventing from the possible insertion of a nominal object between TAM and verb. This analysis allows us to consider PPM as subject affix for VP (in case of pronominal arguments) or subject affix for Aux (in case of a nominal object). It can solve the syntactic puzzle of PPM occurring together with a nominal subject (two syntactic subjects). Considering them as subject affixes instead of the second syntactic subject makes them similar to “standard” 2-stage subject affixes (Creissels 2005) as in Swahili or Latin – the languages with obligatory subject affixes and the so-called pro-drop phenomenon.
always copies the tone of the closest syllable of adjacent NP): a portmanteau pronoun as in (6) for a pronominal object, the last syllable of the nominal object NP (7), or a pronominal subject for intransitive verbs (8).

(7) Gban (personal data)

\[
\begin{array}{llll}
\text{Mú} & \text{ê} & \text{blú} & \text{bé} \\
\text{person} & 3\text{SG.PRS} & \text{bread} & \text{take.PRS}
\end{array}
\]

‘The person takes the bread.’

(8) Gban (personal data)

\[
\begin{array}{ll}
\text{ tô} & \text{tà} \\
1\text{SG.PRS} & \text{go.PRS}
\end{array}
\]

‘I go.’

These examples are interesting in two respects. First, they show tonal regressive assimilation of the verb (assimilating to the tone of the closest syllable), confirming the closer link between object/pronominal subject and the verb than it is attested between the verb and nominal subject NP, etc.). Second, the way the assimilation takes place shows a sort of ergativity in Gban Present tense – the verb always tonally “agrees” with the subject of an intransitive verb or the object of a transitive verb (and never with the subject of a transitive verb).\(^5\)

Summing up, I tried to demonstrate the closer connection between pronominal subject, TAM marker (which is often tonal), pronominal object (all of them forming PPM) and the verb, on the one hand (examples (6) and (8)), and nominal object NP and the verb (7) on the other hand.

The only gap that remains is demonstrating the connection between the PPM complex and the nominal object NP. It can be filled by examples showing the case when a nominal object occurs with a pronominal possessor.

\(^5\) The reasons for this are merely formal and not semantic, so this phenomenon may be called “formal ergativity”.
(9) Gban (personal data)

\[ \text{Má} \quad \text{blú} \quad \text{bé} \]

1SG.SBJ [3SG.POSS.PST bread] take.RECPST

‘[I his] bread took.’

\textit{Má} is a fusion of 3 functions: 1SG.SBJ + 3SG.POSS + PST. Syntactically ‘his’ is a part of the object NP, but morphologically it is a part of the PPM.

We can see that in Gban, the VP forms a certain phonological unity, which seems to have some similarity with the phonological unity of an object and verb in incorporated VP in Tigemaxo (the most phonologically united VP in Mande languages with “real” incorporation). In contrast, incorporated/non-incorporated VP pairing is not found in Gban. The lack of transitive/intransitive formal distinction in the verbs (as attested in Tigemaxo and Soninke) and strict S-O-V order seem to block such variability. The noun of the compound form is still specific/referential, and this is what rules out its characterization as properly “incorporated”.

Another interesting phenomenon found in Gban is the existence of constructions with semantically “empty” verbs, in which the semantics of action is expressed by the object (“action-denoting nominals” (Khachaturyan 2017: 20)), the verb being just a formal predicate forming the so-called “light verb construction”, a phenomenon functionally different from incorporation albeit somewhat similar on the surface:

(10) Gban (personal data)

\[ \text{Ĩ̀ sȍkó bò} \]

1SG run do

‘I run.’

While having an object, the constructions are treated as intransitive. There is no overt marking of intransitivitiy in Gban, but it can be checked in causative constructions. The causative for example (10) is
formed in example (11) with the same marker $k\ddot{a}$ that is used with the intransitive verb ‘to go’ in (12), but a different marker $y\ddot{e}k\ddot{e}$ is used with the transitive verb ‘to take’ in (13).

(11) Gban (Zheltov 2012: 202)

\[
\text{Mà} \quad s\ddot{e}k\ddot{o} \quad b\ddot{o} \quad k\ddot{a}
\]

1SG.SBJ.3SG.OBJ.PRS  run  do  CAUS

‘I make him run.’

(12) Gban (Zheltov 2012: 202)

\[
\text{Mà} \quad t\ddot{a} \quad k\ddot{a}
\]

1SG.SBJ.3SG.OBJ.PRS  go  CAUS

‘I make him go.’

(13) Gban (Zheltov 2012: 202)

\[
\text{Mà} \quad y\ddot{e}k\ddot{e} \quad b\ddot{l}\ddot{u} \quad b\ddot{e}
\]

1SG.SBJ.3SG.OBJ.PRS  CAUS  bread  take

‘I make him take bread.’

Intransitive behavior of the verb with the object looks similar to the incorporated constructions in Tigemaxo and Soninke in a sense that compound object (but in Gban only with “light verbs”) can detransitivize the verb. However, again we cannot find any alternative (referential/non-referential) variants for the same verb.

4. Typological remarks

According to typological theory (§1), the most crucial feature for the language to be treated as “incorporating” is not the phonological incorporation of an object into a verb, but the existence of parallel incorporated/non-incorporated forms and non-referentiality of the incorporated object on the functional side. In “classical” incorporating languages, this feature is combined with detransitivizing of the incorporated verb and phonological compactness of the object-verb.

\[\text{In examples (11)–(13), the lexical tone of the verb is assimilated by the tone of the formal object as it is obligatory for the Present tense in Gban.}\]
complex, the phonological traits not being necessary for counting as “incorporating language” (Mithun 2000: 920–921). Mandinka, Tigemaxo, and Soninke (§2) suit this definition of incorporation but have some peculiarities: in Mandinka, detransitivizing is not attested, while in Tigemaxo and Soninke, it seems to be obligatory.

On the other hand, it is evident that Gban and Gban-like languages (§3) cannot be considered as “incorporating” as they do not have parallel incorporated/non-incorporated forms. At the same time, a close unity of the VP is evident for such languages, and even the syntactic process of detransitivization can be seen in causative constructions. As a summary of the facts, a typological matrix can be presented as follows (Table 1; incorporating languages are marked in grey).

| The typological features of OV incorporation / “formal incorporation” |
|-----------------------------|----------------|----------------|----------------|----------------|
| parallel incorporated / non-incorporated forms | Gban | Koryak, Soninke, Tigemaxo | Mandinka | Russian, Latin |
| detransitivizing | +/- | + | - | - |
| object-verb phonological fusion | + | + | + | - |

At first glance, Gban should be put in the table just next to the left from the column for Russian and Latin. Koryak, Tigemaxo, Soninke, and Mandinka are incorporating languages. Koryak, Soninke, and Tigemaxo being the most consequently incorporating language, could be put to the left. Then Mandinka could go as having the most important feature of incorporation as well as Koryak, Soninke and Tigemaxo, but showing some peculiarities with “secondary” features. Then Gban should be placed as having just secondary features of incorporation, and then
languages like Russian and Latin which do not have any features of incorporation at all should be put. Indeed, both Russian-like and Gban-like languages have no parallel incorporated/non-incorporated forms, while the other languages in the table do have them. However, I would like to put forward the idea that it is Russian and Gban that are the most distant to each other because the nature of having no parallel incorporated/non-incorporated forms is different (or even opposite): Russian really does not have any incorporated forms (the language does not have any formal means for object-verb compounding even in case of pronominal objects), while Gban, on the other hand, does not have parallel “non-incorporated” (in fact, not compound ProSBJ+O+V) forms due to the combination of its typological features (first of all, strict S-O-V order, lack of double object constructions and transitivity/intransitivity markers, a tendency to “glue” PPM-O-V group). Thus, the typological opposition of two elements (incorporating/non-incorporating) turns into the three-element typology: “always incorporating” languages (Gban), “never incorporating” languages (Russian), “incorporating/non-incorporating languages” (Koryak, Soninke and other languages with parallel incorporated/non-incorporated object forms (Table 2).

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tentative typology of OV incorporation / “formal incorporation”</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Attitude to incorporation</strong></td>
</tr>
</tbody>
</table>

This typology can be supplemented by, at least, one more type of relevant but different phenomena presented by Bantu languages. In the

---

7 It can be explained by the very short segmental forms of pronouns and TAM markers (often a vowel or a tone) and the importance of tonal contour.
case of a pronominal object, it is always indexed in the verb, and its functioning as an independent pronoun is ungrammatical, see (14).

(14) Swahili (personal data)
   a. *Ni-na-penda
      1SG.SBJ-PRS-3SG.OBJ-love
      ‘I love him/her.’
   b. *Ni-na-penda ye ye
      1SG.SBJ-PRS-love 3SG
      ‘I love him/her.’

In the case of a nominal object, it is never incorporated but can be indexed (15b) or not indexed (15a) in the verb with the object affix of the corresponding noun class.

(15) Swahili (personal data)
   a. Ni-na-ø-ona
      1SG-PRS-Ø-see 1CL-person
      ‘I see a person / somebody.’
   b. Ni-na-mw-ona
      1SG-PRS-3SG.1CL-see 1CL-person (huyu)
      ‘I see the person.’

These examples, to a certain extent, can be treated as a sort of inversion of “incorporating type”: nominal object can be only indexed in the verb by the affix, and if it is indexed (16) comparing to incorporating languages, the meaning of the object is more referential and specific than when it is not (15). So, the typology turns into a four-element matrix (Table 3).

Indeed, both Gban-like “formal incorporation” and Bantu-like pronominal indexation should be (and are) distinguished from “classical” OV incorporation. However, if we compare incorporation with some different types of O + V compounds, we can make the following interesting observations.

1. In the traditional opposition between incorporating and non-incorporating languages, the latter show some essential differences: some of them never have any VO compounds (Russian), while others (Gban) always have VO compounds and hence take distant rather than neighboring positions in typological scale.
2. The languages with a verbal indexation are different from both incorporating and non-incorporating language, but share with the former ones one important common feature: in the case of a nominal object, they can have parallel (with pronominal indexation / without pronominal indexation) constructions in which the indexed one point to (more) referential status of an object.

3. On the other hand, the languages with verbal indexation share another important common feature with the languages with “formal incorporation”.

Table 3

<table>
<thead>
<tr>
<th>Attitude to incorporation</th>
<th>“Always incorporating” (no parallel not compound OV forms (“formal incorporation”))</th>
<th>With parallel incorporated / non-incorporated (indexed / non-indexed) object forms</th>
<th>“Never incorporating” (no object incorporation/compounding/indexing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always indexing (for pronominal arguments only), referential (for indexed nominal objects) / non-referential (for non-indexed nominal objects)</td>
<td>Intransitive/non-referential (for incorporated objects), transitive/referential (for non-incorporated objects) = “classical incorporation”</td>
<td>Language</td>
<td>Gban</td>
</tr>
</tbody>
</table>
incorporation” (obligatory OV compounding): in the case of a pronominal object, the verbal indexation (pronominal “incorporation”) is obligatory in a similar way as object compounding is obligatory in Gban-like languages.

4. Hence, we can see a sort of typological scale: Gban-like languages > Swahili-like languages > Koryak-like languages > Russian-like languages. The left position is occupied by the languages with obligatory OV compactness, then go the languages with verbal indexation (indexation obligatory for a pronominal object, optional and functionally motivated – for nominal objects), then – incorporating languages (incorporation always optional and functionally motivated), then – the languages without any OV compactness.

The author realizes quite well that for proper typology, a representative set of languages should be used. In this case, we have a “typology of phenomena” rather than a typology of languages. It is very likely that the list of relevant phenomena can also be expanded. However, at the moment, the author tried to collect different types of OV relations and formal markings, and the purpose of the paper was to show that the traditional binary oppositions (incorporating/non-incorporating, indexing/non-indexing) can form more gradual, complex, and, hence, interesting picture if to analyze them in the complex. The Mande languages which are the focus of this paper give very interesting data of both incorporation and related phenomena.

Acknowledgements

This paper, as well as the whole issue of the journal, is devoted to Valentin Vydrin. For the author, Valentin is a significant person – friend, colleague, and teacher at the same time. Due to Valentin, the author (as well as many Russian linguists) got his first field experience in Africa. This text was encouraged by the continuing discussion with Valentin about the dynamics of typological attribution of Mande languages. The author is also very grateful to two anonymous reviewers for their valuable critical remarks and corrections of the text.
Abbreviations

- ACP – accomplished
- CAUS – causative
- DEF – definite
- DETR – detransitivization marker
- EP – epenthetic
- FV – final vowel
- GER – gerundive
- IPFV – imperfective
- INTR – intransitivity marker
- NP – noun phrase
- NSBJ – non-subject
- O/OBJ – object
- PL – plural
- POSS – possessive
- PRO – pronoun
- PRS – present tense
- PST – past tense
- RECPST – recent past
- S/SBJ – subject
- SG – singular
- TAM – tense-aspect-modality marker
- TR – transitivity marker
- V – verb
- VP – verb phrase

References


