Abstract

Swedish


Nyckelord: lingvistik, litet korpus, pidgin, tiraljör, andra världskriget, västafrika, senegal, petit nègre, moi-ya-dit, anonym manual 1916, Charles Mangin, Lucie Cousturier, kolonialism

English

Français Tirailleur (FT) is a pidgin language that was spoken by West African soldiers and their white officers in the French colonial army approximately 1857-1954. The aim of this study is to investigate a corpus of previously unanalyzed utterances of FT in order to discern linguistic structures and test previous statements about the nature of FT. Much of previous literature on FT is based on an anonymous manual published by the French military in 1916, this thesis aims to provide new information to our understanding of this pidgin. These are some of the findings: standard negation is expressed by means of a preverbal particle (pas), polar interrogation by intonation, grammatical gender is not a productive category and attributive possession is expressed by possessive pronouns, juxtaposition (possessum - possessor) and prepositional constructions. The standardized type-token-ratio of this corpus, 26%, suggests that the lexicon of pidgins needs to be further studied. Comparisons with corpuses of spoken language are needed. There are two very frequent pre-predicate markers that are considered characteristic of FT: ya and yena. These two markers have previously been described as stative verbs, relativizers and markers of finiteness. The two markers are very frequent in a majority of the sources and are highly polysemous, functioning as stative verbs, copula or copula-like markers and possibly also predicate markers. The status of adjectives as a part-of-speech in FT is also discussed.

Keywords: Linguistics, Corpus, Pidgin, Tirailleur, Sénégalaïs, West Africa, Petit-Nègre, Moi-Ya-Dit, Charles Mangin, Lucie Cousturier, Dakar 1944, First World War, WW1, Second World War, WW2, APiCS, Senegal, colonialism
Contents

1 Introduction 1

2 Research questions 2

3 Background 3
   3.1 Theoretical background ........................................... 3
      3.1.1 Theory, assumptions and frameworks ......................... 3
      3.1.2 Pidgins and contact languages ............................... 5
      3.1.3 Summary of theoretical background .......................... 7
   3.2 Les Tirailleurs Sénégalais, West Africa and France ............ 7
      3.2.1 Brief account of the history of contact between France and West Africa 8
      3.2.2 The pidgin of Français Tirailleur ............................ 11
      3.2.3 Summary of previous statements about the grammar of FT .... 12
      3.2.4 The linguistic makeup of the Français Tirailleur troops ........ 16
      3.2.5 Summary: Les Tirailleurs Sénégalais, West Africa and France ... 18

4 Method 20
   4.1 Source material .................................................. 20
   4.2 Reliability of the sources ....................................... 23
   4.3 Comparison with other languages .................................. 23
   4.4 Limitations of the study – the nature of corpus ................ 23
   4.5 Annotation of the data ............................................ 24

5 Results and analysis 26
   5.1 Data makeup ..................................................... 26
      5.1.1 Clause-types and stative/dynamic ............................. 26
      5.1.2 Type-token-ratio in the corpus ............................... 27
      5.1.3 Summary: data makeup ......................................... 31
   5.2 Function seeking form ............................................ 31
      5.2.1 Polar Interrogative .......................................... 31
      5.2.2 Standard Negation ............................................ 34
      5.2.3 Non-standard negation (jamais, rien, personne, plus) ...... 37
      5.2.4 Grammatical gender .......................................... 38
      5.2.5 Personal pronouns ............................................. 40
      5.2.6 Attributive possession ....................................... 43
      5.2.7 Predicative possession ....................................... 45
      5.2.8 Copula relations ............................................... 46
      5.2.9 ‘Being able’ — moyen and pouvoir ............................ 47
      5.2.10 Summary: function seeking form ............................ 48
   5.3 Form seeking function ............................................. 48
      5.3.1 Drop of overt subject ......................................... 48
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.2</td>
<td>Polysemous form <em>ya</em></td>
<td>51</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Polysemous form <em>yena</em></td>
<td>55</td>
</tr>
<tr>
<td>5.3.4</td>
<td><em>Gagner</em> — to get, become or have?</td>
<td>56</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Wanting, liking and being happy — <em>content</em> and <em>vouloir</em></td>
<td>57</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Reduplication</td>
<td>57</td>
</tr>
<tr>
<td>5.3.7</td>
<td>Summary: form seeking function</td>
<td>59</td>
</tr>
<tr>
<td>5.4</td>
<td>Notes on diachrony</td>
<td>60</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Order of modifier and head</td>
<td>60</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Aspectual <em>fini</em>?</td>
<td>60</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Pre-verbal <em>qui/que</em></td>
<td>60</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Summary: notes on diachrony</td>
<td>61</td>
</tr>
</tbody>
</table>

6 Discussion                                                                 62

7 Conclusions                                                               64
List of Figures

1. Map of European expansion in Africa in 1914 (Guillaume Balavoine 2013) after the *Scramble for Africa* ©2002 Guillaume Balavoine.
6. Type-frequency-ratio of the FT-corpus. Non-logarithmic, zoomed in at first 100 types.
7. Type-frequency of the FT-corpus. X-axis = types ranked by freq, Y-axis=frequency (log-log=10).

List of Tables

1. Brief historical overview of important events in the history of *Les Tirailleur Sénégalais*.
2. Languages proposed as involved in the multilingual situation of FT.
3. FT material sorted by year of production of utterance.
4. The eight largest documents of the corpus.
5. Independent/subordinate clauses.
6. Declarative/interrogative/imperative/other.
7. Stative and dynamic clauses.
8. Types/tokens in FT.
11. Expression of Standard Negation in FT.
12. Expression of non-standard Negation in FT.
13. Gender marking in FT.
15. Pronouns in FT.
16. French possessive pronouns.
17. Expression of attributive possession in FT.
18. Predicate possession in FT.
19. Copula relations distinguished in this study.
20. Expression of copula relations in FT.
21. Drop of overt subject.
22. Null subject distributed over person and number.
23. *Ya* in FT.
List of glosses

1sg 1st person singular
2sg 2nd person singular
3sg 3rd person singular
1pl 1st person plural
2pl 2nd person plural
3pl 3rd person plural
ACC Accusative case
AP Adjective Phrase
COP Copula
DAT Dative case
DEF Definite article
DEM Demonstrative
DN Double Negation
FUT Future tense
INDEF Indefinite article
NEG Negation
NOM Nominative case
NP Nominal Phrase
PASS Passive
POSS Possessive
PP Prepositional Phrase
PROG Progressive aspect
PRS Present tense
Pro Pronoun
PST Past tense (“Simple” Past)
PTCP Participle
SN Standard Negation
Ya the auxiliary/stative verb/predicate marker ya.
Yena the auxiliary/stative verb/predicate marker yena
List of abbreviations

APiCS  Atlas of Pidgin and Creole Structures (see Michaelis et al. 2013)
BLT   Basic Linguistic Theory (see Dixon 2010a)
ELICOP Étude Linguistique de la Communication Parlée (see Debrock et al. 2001)
FT    Français Tirailleur
TTR   Type-token-ratio. The number of types divided by the number of tokens, usually displayed in percent.
sTTR  Standardized type-token-ratio (only of 2 000 consecutive words)²
SIL   Summer Institute of Linguistics - faith-based nonprofit organization founded in 1934. SIL is devoted to the description, documentation and preservation of the world’s languages.
stdzd Standardized material
WALS  Word Atlas of Language Structures (see Dryer and Haspelmath 2011)
WW1   First World War
WW2   Second World War

Format of examples from corpus

Each example from the corpus contains a reference-ID that identifies the entry in the corpus (ex. FrTir0286b) and a reference to the original source. Please note that the date of utterance is not necessarily the same as the date of publication of document. The examples are given in four lines.

Original       The string exactly as in the source
Stdzd          The string converted to a more standardized spelling as to make the corpus more coherent, comparable and searchable (by Mikael Parkvall)
Gloss          Glossing of the stdzd-string (by Mikael Parkvall)
Translation    Translation into English, either by the source or Mikael Parkvall

Example:
Français Tirailleur (FrTir0284, Anon 1916:54)

Original  Ramener jambe droite en avant, genou droite bien fléchi
Stdzd     Ramener jambe droite en avant, genou droite bien fléchi
gloss     bring leg right forward, knee right well/very bent

'Put your right leg forward, with your knee bent’
1 Introduction

This thesis is a descriptive study of a contact language known as Français Tirailleur (FT) based on corpus material.

When speakers of different languages have to communicate despite a lack of common language several things can happen; they can settle on one of the involved languages as a common tongue, but they can also create a new language variety. These languages are called contact languages, and this thesis is devoted to the description of one such language variety, FT.

FT is a contact language with no mother tongue speakers — a pidgin. It was spoken by West African soldiers and their white officers, approximately between 1857-1954. These West African soldiers were enrolled in the French Colonial army and had to communicate, not only with superior white French officers but also with the fellow soldiers with whom they did not necessarily share a common language. The sentence in (1) was uttered by a West African soldier, a Tirailleur Sénégalais, in the 1940’s and illustrates the situation quite well.

(1) Français Tirailleur (FrTir0377, Woodfork 2001:108)

Original  Oui, parce que lui parler sa langue, je comprends pas, et moi parler ma langue, lui comprend pas

Stdzd Oui, parce que lui parler sa langue, je comprends pas, et

Gloss yes, because 3sg speak 3sg.POSS language 1sg understand NEG, and

1sg speak 1sg.POSS language, 3sg understand NEG

‘Yes, because if he speaks his language, I don’t understand, and if I speak mine, he won’t understand’

There are several different types of contact language varieties and there are also many different definitions of these varieties. As Muysken and Smith (1995:3) writes “[c]reolists agree neither about the precise definition of the terms pidgin and creole, nor about the status of a number of languages that have been claimed to be pidgins and creoles”. One definition of a contact language that is often cited is that by Thomason (1997:3): “a language that arises as a direct result of language contact and that comprises linguistic material which cannot be traced back primarily to a single source language”.

While there are many different theories concerning the nature and origin of contact languages, one distinction that many scholars do agree on is that there are contact language varieties that have native speakers and there are those that do not. This thesis is concerned with the latter and specifically languages that are stable enough to be labeled ‘pidgins’. The subject matter of what is and what is not a creole and how they came about, while highly interesting and worthy of more debate, is not discussed in this thesis.

Contact languages usually include elements from many different languages, but the bulk
of the vocabulary is often from one language — this is what is known as the lexifier (or superstrate) language. Most languages of the world contain words and elements from other languages, what makes pidgins in particular different is not the mix of linguistic material *per se*, but rather the reduction of grammar and lexicon. Pidgins have been noted to arise in situations of great need such as war and slave trade but also through trade (Foley 2006).

The language variety that we are concerned with in this thesis goes by several names, among others: *Forofifon Naspa, Moi- ya-dit, Petit-Nègre* and *Français Tirailleur*. I will in this thesis use the term *Français Tirailleur* (FT). The word *tirailleur* in French translates to infantryman or rifleman in English. African soldiers from Algeria and Morocco were also called ‘Tirailleurs’, this thesis will however only concern Tirailleurs of West Africa, *Les Tirailleur Sénégalais*.

There is one source on FT that has been cited more often than any other, the anonymous manual from 1916 (*Le français tel que le parlent nos tirailleurs Sénégalais*). This thesis aims to provide more insights into the nature of FT, using not only material from the well-cited anonymous manual but from many more documents and authors.

The corpus of FT consists of material from 148 different documents, 130 authors, 8686 words and 1737 strings. This material has been collected, glossed and translated by contact language researcher Mikael Parkvall and annotated by me.

### 2 Research questions

The aim of this thesis is to investigate the structure of Français Tirailleur (FT) through a small corpus. The major questions are

i. What grammatical features can we discern from this corpus?

ii. How does FT differ from the lexifier French with respect to those features?

iii. Is there any evidence of influence from West African languages with respect to those features?

iv. Is there evidence of change over time?

Much of the literature on FT has focused on the anonymous manual from 1916, one of the aims of this thesis is to test if there is support for the statements that have been made previously in earlier literature.
3 Background

The background section is divided into two parts. The first part concerns the theoretical background of the study and the second part provides a background for the multilingual situation where this variety arose.

3.1 Theoretical background

3.1.1 Theory, assumptions and frameworks

A language variety is a conventionalized system for expressing meaning between humans. Languages are in their nature fluid, both synchronically and diachronically. This means that the task of a linguist (understanding human language) is very hard indeed. There is a fine line between what we define as two separate languages and what we label as dialects of one language. Contact language varieties are by no means not exempt from this. What is called one language variety can, usually, also be described as a collection different varieties, perhaps overlapping in certain areas more than others. This is the reality of a linguist’s object of study and something we always need to keep in mind.

A problem of research on pidgins in particular is that the object of study is more unstable in their structure than languages on average. It is necessary to take this variation into account and recognize the limitations of what we can understand and describe in a language. This is the reason why the focus of this study is on basic structures which we expect to be more stable in this pidgin variety and more frequent in our corpus, such as negation, polar interrogation and attributive possession.

Theoretical linguistics have suggested that this variation and change is not random, there are patterns in this variation; some features might be dependent on others and certain paths of language change might be more common than others.

I will in this thesis make use of the theoretical concepts, assumptions and framework(s) as found in much of descriptive linguistics and linguistic typology, such as WALS (Dryer and Haspelmath 2011). Many general linguists might work outside specific theories and frameworks, such as Minimalism or Role and Reference Grammar, but there are nonetheless still certain theoretical assumptions present.

Many of the assumptions that are widely used in language description and linguistic typology are summarized in the works on Basic Linguistic Theory (BLT) by R.M.W Dixon and Matthew Dryer. BLT aims at providing a cumulative framework of descriptive linguistics, not an explanatory one. BLT contains many concepts that are widespread and used by many different scholars in the field of theoretical linguistics.

Ideally, descriptive linguists and typologist should work with descriptions of language that are the result of an analysis of those language on their own terms and not based on knowledge of categories in other languages. This is what Haspelmath (2010b) calls “framework-free”; descriptive linguists should employ a set of concepts that are constructed separately for each language. This is, sadly, seldom feasible, and in the case of FT most likely impossible. We cannot gather new data, and what is more relevant: we have little
or no negative evidence illustrating what is ungrammatical.

How are we then to do cross-linguistic comparisons at all? Linguists use labels that are intended to cover a certain shared function/meaning — cf. Haspelmath’s Comparative Concepts, the gram-types of the ‘Bybee&Dahl-approach’ (Bybee and Dahl 1989) or the labels of BLT. These concepts are useful for linguists as they facilitate communication between scholars and cross-linguistic comparisons, but we should always keep in mind that this does not mean that these categories are the same across all languages.

Many scholars have emphasized the need to distinguish between language-specific descriptive categories and comparative categories intended for cross-linguistic comparison. Some researchers, such as (Haspelmath 2010a), have proposed that these comparative categories are created for the benefit of linguistic typology alone and need not have any bearings on reality. Others, such as Dahl (p.c.), suggest that these cross-linguistic categories (‘gram-types’) might represent probabilistic clusters of linguistic function that have potential bearing on the reality of speakers’ minds.

FT is most likely an unstable language variety or collection of varieties; as are most (if not all) pidgins. FT is probably also a dead language variety (though this needs to be made certain) and the written records are very limited. The modest aim of this thesis is to discern what structure is possible to grasp from the corpus material, focusing on high-frequent phenomena and grammatical structures that are believed to be more stable, such as negation, polar interrogation, possession and the expression of a formal subject.

In order to investigate these features, I will employ concepts and labels that are widespread in language description and in cross-linguistic comparison. The alternative would be to construct new categories based on FT alone. It is, unfortunately, hard to conceive how that could be accomplished using the limited material available.

Furthermore, I make the theoretical assumption that there is such a thing as a conventionalized contact variety (or collection of varieties with considerable overlap in certain areas) with no mother tongue speakers. If it is stable enough I call it a ‘pidgin’ (as opposed to the more unstable kind that we call ‘jargon’). These language varieties are likely to exhibit great internal variation. Whether or not creoles originate from these varieties or not is not discussed here.
3.1.2 Pidgins and contact languages

There are several potential outcomes of situations when speakers lacking a common language need to communicate. These are some of the outcomes linguists have found.

**Language shift** One or more groups adopt the language of one of the groups. This might lead to language death for the other languages involved.

**Multilingualism** Speakers learn the languages of the other group(s) while maintaining their native languages, perhaps using the different language for different functional domains.

**Pidgin** A language variety with no mother tongue speakers and restricted grammar and lexicon, primarily used as a between-groups language.

**Creole** Contact language variety with mother tongue speakers. There is great controversy as to the origin of creoles, that topic will not be discussed here.

**Mixed language** Mix of two (or more) languages resulting in a variety that shows positive genetic relationship with more than one language (Bakker 1997:195). Not to be confused with creoles, jargons or pidgins.

A pidgin is a type of contact language with no native speakers. Pidgins are characterized by reduced grammar and lexicon. Thomason (1997:76) writes that a pidgin prototypically “arises in a new contact situation in which three or more groups of speakers come together for purposes of trade or other limited communicative purposes”.

Thomason’s definition excludes a purely bilingual contact situation, according to her it is the three-language-scenario or more that is key. There are counter examples of this, such as Russenorsk (Broch and Jahr 1981) — a pidgin formed in a bilingual contact situation of Russian and Norwegian. In the case of FT, however, there is no doubt that the West African group spoke several different languages resulting in a multilingual situation of at least more than three languages.

In his book *La Force Noir* Mangin (1910) writes that the largest ethnical groups among the French West African colonial troops were Wolof, Fula 4, Hausa and Mande, see section 3.2.4 for more details. These groups of native West Africans did not seem to know each other’s languages very well, nor did they, typically, master French — the language spoken by their commanding officers 5. This means that FT was not only necessary for communication between the French officers and the West African soldiers but also among the soldiers themselves (see example 1 on page 1). It is not likely though that FT was used between

---

3 Mixed languages typically show similarities with two (or more) in their grammar as well as their lexicon, a language that draws only lexicon from one language and only grammar from another is usually not labeled “mixed”

4 This group is also known as Pulaar, Fulfulde and Peul.

5 Though there are a few exceptions of West Africans who did mastered French.
speakers in these troops that did know each others language, FT was in other word a between-groups-language rather than a in-group-language. Pidgins tend to be quite unstable and there are very few descriptive works of pidgins. We do not know much about pidgins and it is hard to say how many pidgins there are or have been. Pidgins are not clearly defined, where do we draw the line between interlanguage, foreigner talk, jargon, pidgins and creoles?

Holm made an important contribution to the field of contact linguistics with his two-volume survey of pidgins, semi-creoles and creoles that was published in 1988. The two volumes contain 88 contact languages. Ethnologue - catalogue of the world’s languages distributed by SIL - Lewis et al. (2013) contains 17 pidgins (and 93 creoles). Glottolog (Nordhoff et al. 2012) another catalogue over language varieties, contains 35 pidgins. The same site that hosts Glottolog also contains Langdoc, a catalogue of language descriptions. Langdoc contains 51 entries of grammars or grammar sketches of pidgins.

The World Atlas of Language Structures (WALS) and The Atlas of Pidgin and Creole Language Structures (APiCS) do not make an explicit distinction between pidgins and creoles (or any other type of contact language). WALS contains 32 pidgins and creoles (Dryer 2011b) and APiCS encompasses 76 pidgins and creoles (Michaelis et al. 2013).

Pidgins emerge from language contact, but if the groups have too much contact the speakers might switch to one of the involved languages or the pidgin might expand into a creole (depending on your analysis of the origin of creoles). Holm (1988:4-5) includes this in his definition of a pidgin “no groups learns the native language of any other group for social reasons that may include lack of trust or of close contact”. However, if the groups have too little contact the language variety might be less stable and consistent — not a distinct language variety at all.

More than 230 language varieties have at least once been defined as pidgins or potential pidgins by expert scholars (Parkvall p.c.). Of the language varieties that are included in WALS and APiCS there are 6 and 16 respectively that have at least once been labeled “pidgin”. It is important to note that neither WALS nor APiCS claim to include all pidgins or creoles.

Why are there then so few descriptions of pidgins? There are many possible reasons

**Not enough research, yet** The study of contact language varieties is itself a very young discipline

**Too little data** There are very few documented pidgin utterances seeing as it is not a language variety that is often expressed in writing.

**Lack of knowledge** Speakers of pidgins are not always aware of the fact that they speak a distinct language variety

**Shame** Even if they are aware that they master a pidgin language variety it might be connected to social stigma — motivation from the speakers in describing the variety is very low
Lack of in-group identity  Connected to the two previous points; pidgins have no native
speakers. The speakers all belong to other groups and are perhaps not likely to see
this group in which the pidgin arose as important enough compared to their native
group, again making motivation from the speakers in describing the variety is very
low.

Many of the problems in pidgin studies are similar to problems faced by scholars studying critically endangered or dead languages.

As previously stated, we as linguists know little about the nature of pidgins, but what we do know is that they seem to arise in situations of attenuated contact such as colonialism, slave trade, war or trade (Foley 2006:2). There is no time to learn each other’s languages thoroughly and the need to communicate seems to be stronger than the desire to belong to a certain group\textsuperscript{6}, the motivation behind pidgins is perhaps what sets them apart from second-language-acquisition — further research on these matters is however needed.

3.1.3 Summary of theoretical background

Ideally, one would describe each language on its own terms — using bottom-up generated
categories. This is however not feasible in the case of FT, this thesis will instead make
use of concepts and categorizes that have been used in linguistic typology and language
description.

Pidgins are contact-varieties that are characterized by reduced grammar and lexicon
and the lack of mother tongue speakers. Pidgins are shaped by a multilingual situation
where several groups need to communicate, but lack a common language. These situations
need not produce pidgins, there are other possible outcomes. We know little of the nature
of pidgins.

3.2 Les Tirailleurs Sénégalais, West Africa and France

This section contains a background of the contact between France and West Africa and of
the West African troops, Les Tirailleur Sénégalais.

The first West African soldiers were enrolled in the French colonial army in 1820 and
the company of the Tirailleur Sénégalais was formed in 1857 (Echenberg 1986:311-315).
They fought for France in both World Wars. The last company of West African soldiers in
the French army was disbanded in 1964 and the last Tirailleur who served in WW1 died

It is important to keep in mind that while slavery was abolished in France and her
colonies in 1848, this does not mean situation changed totally over night. There is an
inter-departmental report from 1950 suggesting that the recruitment was not necessary
voluntary from that point forward. The French military was to go “up-country to enroll
captives, to whom the sum needed to purchase their freedom is given as enrollment bounty”

\textsuperscript{6}The group being native or native-like speakers.
In other words, they more or less bought captives and turned them into soldiers.

### 3.2.1 Brief account of the history of contact between France and West Africa

There is evidence of French presence in West Africa since the 16th century. The French colonized West Africa, extracting slaves, gold and other goods. The English, Spanish, Portuguese and to a less extent the German, Danish and Dutch were also present in the area.

On the following page is a map (figure 1) illustrating the entire African continent of 1914 (Guillaume Balavoine 2013). This map shows the partition of Africa after the *Scramble for Africa* of 1881-1914. The *Scramble of Africa* is a name given to a intense period of invasion, colonization and annexation of the africa continent by European colonial powers.

France was at that time in possession of great parts of West Africa (*Afrique Occidentale Française*), Equatorial Africa and the entire island of Madagascar. Liberia and Abyssinia (present day northern Ethiopia and Eritrea) were independent. The states of West Africa were all independent by the end of the 1960’s (Chafer 2002).

At the time of the French colonization of West Africa many Africans were labelled as “Sénégalais” even if they were not from Senegal. The is also reflected in the name of the West African troops in the French colonial army; *Les Tirailleur Sénégalais*.

Table 1 contains a brief overview of some of the major events in the history of the contact between France and West Africa. For a more extensive account, see (Wilson 1999:3-21).
Figure 1: Map of European expansion in Africa in 1914 (Guillaume Balavoine 2013) after the *Scramble for Africa* ©2002 Guillaume Balavoine.
<table>
<thead>
<tr>
<th>Time Period</th>
<th>Event Description</th>
<th>Source/Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1340-1470</td>
<td>European expansion along the African coast</td>
<td>Thornton (1992:32)</td>
</tr>
<tr>
<td>1444</td>
<td>The Portuguese start extracting gold and slaves from Senegal</td>
<td>Thornton (1992:32)</td>
</tr>
<tr>
<td>1628</td>
<td>Unsuccessful attempt by the French to settle permanently in Senegal</td>
<td>Biondi (1987:37)</td>
</tr>
<tr>
<td>1635</td>
<td>“The population of Rufisque² speak a fairly intelligible kind of French [...] they pronounce in our language offences and swearwords”</td>
<td>Delafosse 1931:11</td>
</tr>
<tr>
<td>1658</td>
<td>St-Louis-du-Sénégal is established</td>
<td>Biondi (1987:37)</td>
</tr>
<tr>
<td>1817</td>
<td>Slave trade from St-Louis stopped, the export shifted to gum</td>
<td>Wilson (1999:10)</td>
</tr>
<tr>
<td>1820</td>
<td>First African soldiers in the French colonial army</td>
<td>Echenberg (1986:311)</td>
</tr>
<tr>
<td>1823</td>
<td>First all-African company in the French colonial army</td>
<td>Echenberg (1986:312)</td>
</tr>
<tr>
<td>1848</td>
<td>Slavery is abolished in France and her colonies</td>
<td></td>
</tr>
<tr>
<td>1857</td>
<td>Official creation of <em>Tirailleur Sénégalais</em></td>
<td>Echenberg (1986:315)</td>
</tr>
<tr>
<td>1910</td>
<td>Charles Mangin propagates for the use of African troops in a potential European war</td>
<td>Mangin (1910)</td>
</tr>
<tr>
<td>1914</td>
<td>WW1 starts</td>
<td></td>
</tr>
<tr>
<td>1916</td>
<td>The anonymous manual of FT is published</td>
<td>Anon (1916)</td>
</tr>
<tr>
<td>1916</td>
<td>WW1 ends</td>
<td></td>
</tr>
<tr>
<td>1939</td>
<td>WW2 starts</td>
<td></td>
</tr>
<tr>
<td>1944</td>
<td>24-35 West African soldiers are killed by French officers in Dakar, Senegal, following a mutiny³</td>
<td>Echenberg (1978)</td>
</tr>
<tr>
<td>1945</td>
<td>WW2 ends</td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td>All states of West Africa are independent</td>
<td>Chafer (2002)</td>
</tr>
</tbody>
</table>

Table 1: Brief historical overview of important events in the history of *Les Tirailleur Sénégalais*

---

²Rufisque is a coastal town in Senegal. It was an important town in its own right, but has now become a suburb of Dakar, the capital. In Wolof the town is called Tëngëj.

³The movie *Camp de Thiaroye* by Ousmane Sembène depicts this story.
3.2.2 The pidgin of *Français Tirailleur*

The language of the West African soldiers in the French colonial army has been mentioned in descriptive works from the 19th century and forward. The earliest documented utterances in FT are found in Dupratz (1864).

(2) Français Tirailleur (FrTir1234, Dupratz 1864:398)

**Original** Mission n’y a pas bon; toujours en la classe, toujours en l’étude; non mission n’y a pas bon

**Stdzd** Mission n ya pas bon; toujours en la classe, toujours en l’-étude; non mission n ya pas bon

**gloss** mission NEG1 ya NEG2 good; always/still in DEF class, always/still in l’-étude; non mission n ya pas bon

DEF-class no mission NEG1 ya NEG2 good

‘The mission is no good. There’s too much studying. No the mission sure is no good’

Maurice Delafosse wrote about FT in 1904, describing it as a French equivalent to the more well-known English pidgins of the area.

The most cited source on the language variety is an anonymous manual, *Le français tel que le parlent nos tirailleurs Sénégalais* (see figure 2). The manual was printed in 1916 and was intended to facilitate the communication between French officers and the African soldiers in the French army.

![Figure 2: Cover of the anonymous manual of FT. ©1916 Imprimerie Militaire Universelle](image-url)
The manual is prescriptive, informing white officers how they should best formulate orders for optimal effect. The author(s) does make comments that suggest that the material is based on some actual experience with West African soldiers. There is also references to the structure of Bambara that implies that the FT found the anonymous manual is the product of a conscious effort rather than natural utterances.

The manual is divided into two parts, a description of the language variety (here labeled ‘DESC’) and a set of commands and phrases (‘COMM’). Wilson (1999) has analyzed the manual extensively and she suggests that it is possible that the two parts were written by two different people. The material in the anonymous manual concern the life of the soldier and war, there is little mention of anything not related to the military. (3) is an typical example of a sentence from the anonymous manual:

(3) Français Tirailleur (FrTir0257a, Anon 1916)

Original

Si ennemi ya gagné blessés trop, tués trop

Stdzd Si ennemi ya gagner blesser trop, tuer trop

gloss if enemy ya have/get wound/wounded much, kill/killed much

‘If the enemy has many wounded and killed’

Delafosse (1904) and the anonymous manual of 1916 do not only contain utterances in FT but also statements about the grammar of FT. The manual has been analyzed by Wilson (1999) and she provides further insights into FT based on the manual. There is also a brief description of FT in Corne (1999). The next section of this thesis contains a summary of the statements that have been made about the linguistic structure of FT.

3.2.3 Summary of previous statements about the grammar of FT

This section contains a list summarizing the statements that have been made about the structure of FT Delafosse (1904), Anon (1916) and Wilson (1999). The statements from Anon (1916) are also accompanied by page reference to the thesis by Wilson as it contain translations into English and a valuable commentary.

Anon (1916) often gives para-linguistic advice to French officers on how they should communicate with the black soldiers, for example Désigner toujours le même objet ou exprimer la même idée par le même mot [translation into English by Wilson: always use the same word to express same object or ideal]. Statements about FT in this manner are not included in here, this summary contains explicit statements about the linguistic structure of FT.

I have aimed at summarizing the statements in a clear and concise manner. As said before, there is variation in FT and there are statements in this summary that will contradict each other. I will first give a summary of statements on FT made by previous literature that concerns this investigation directly. Following that there is a summary of statements
that have been made about FT in the literature, but that will not be discussed directly in this study. The list of statements is structured in the same way as the result section of this thesis.

**Polar interrogation.** Anon (1916:16) and Wilson (1999:47) claim that polar interrogation in FT is expressed by intonation alone. This is addressed in section 5.2.1 of this thesis.

**Standard/verbal Negation** (SN). Wilson (1999:92) claims that SN is expressed by pre-verbal *pas* alone, but follows, what she labels, the finiteness marker and relativizers *ya* and *yena*. (Delafosse 1904:265) claims that *pas* is positioned after the verb. Section 5.2.2 of this thesis is devoted to the expression of SN.

**Grammatical gender.** Anon (1916:7) claim that all inanimate objects are assigned the etymologically masculine gender (*le, un, mon*). This can be interpreted as statement of FT lacking a distinction in grammatical gender. Furthermore, the anonymous manual claims that the natural gender of animate nouns is expressed by the suffix *femme*, i.e. *chien-femme* means 'bitch'. Section 5.2.4 of this thesis is devoted to the expression of gender in FT.

Anon (1916:7) and Delafosse (1904:265) says that there are no articles (indefinite or definite). There are however cases where the indefinite article is perceived as belonging to the word, for example: *mon latête*. The issue of agglutinating articles and possessive pronouns is also dealt with in section 5.2.4.

**Personal pronouns.** Anon (1916) claims that there is no distinction between subject and object in the personal pronoun system. FT doesn’t have any special forms for reflexive pronouns either (Wilson 1999:92). The pronouns are: *moi, toi, lui, nous, vous* and *eux* (Anon 1916:11). An overview of the personal pronouns of the FT-corpus is found in 5.2.5.

**Attributive possession.** Anon (1916:9), Wilson (1999:34) and Delafosse (1904:265) gives several strategies in FT for the expression of attributive possession (constructions where the possessum and the possessor form a noun phrase).

When the possessor is a 1st or 2nd person singular pronoun FT is said to employ possessive pronouns (*mon, ton*). For the other persons and nominal possessors FT has a prepositional construction with *pour* ‘for’. In the *pour*-construction the order is possessum - *pour* - possessor. Delafosse (1904:265) also describes a prepositional construction, but with *de* ‘of’ instead of *pour* ‘for’.

Attributive possession can also be expressed through juxtaposition (no prepositions, no possessive pronouns). The anonymous manual says that the preferred order is possessor - possessum, but the opposite is possible as well. Example: *tirailleur fusil*, meaning ‘the soldier’s rifle’ [sic] (Anon 1916:14 and Wilson 1999:43, 84).

It is noted by the anonymous author(s) that the order possessor - possessum is most likely a result of influence from substrate languages. If a speaker uses possessum - possessor, the anonymous manual (page 15) says that “ils seront obligés de faire mentalement l’inversion pour rétablir l’ordre qui, pour leur langue, est l’ordre naturel” [Eng: ‘they will have to mentally reverse the order so as to retain what is for them the natural order’]. In

---

7Possessum: that which is possessed, also known as ‘possessee’.
other words, to facilitate communication with the soldiers the officers are encouraged to use structures the author(s) believe are present in the substrate languages.

The expression of attributive possession is found in section 5.2.6 of this thesis.

**Predicative possession.** Anon (1916:13) and Wilson (1999:44) states that the most frequent expression of predicative possession is a construction involving *ya* and *gagné*. Predicative possession is discussed in section 5.2.7 of this thesis.

**Copula relations and adjectives.** Copula is a linguistic term that can have many different meanings in the literature. Wilson (1999:81) Corne (1999:2000) says that FT has no copula, and I believe they are referring to and equative, predicative and/or identity-copula. The anonymous manual notes that the French copula-verb *être* ‘be’ doesn’t exist in FT and that it is sometimes replaces by *ya* (Anon 1916:13). Wilson writes that stative verbs *ya* and *yena* can be used in similar contexts as predicative copulas. This has implication for the interpretation of adjectives as a part-of-speech in FT. Copula relations are found in section 5.2.8 and adjectives are discussed in section 6.

Anon (1916:8-9) claims that adjectives in FT are not expressed by a modifying word within the noun phrase but rather with a relative clause containing a *ya* or *yena* and a 3rd singular pronoun. Anon illustrates this with the example that *un enfant bon* [Eng: a good child] becomes *un enfant (que) il est bon* [Eng: a child (that) he is good] (Anon 1916:8-9 and Wilson 1999:33). It is not clear if these are to be seen as finite clauses or not.

Wilson (1999:81) states, after analyzing the phrases listed by the anonymous manual, that there are adjectives expressed in the same manner as in the lexifier and that they can be pre-posed or post-posed what the modify.

**Moyen.** (Wilson 1999:52) notes that the notion of ‘be able to’ is *moyen* (as opposed to French *pouvoir* ‘be able to’), this is investigated in section 5.2.9.

**Ya.** One of the most frequent items in the corpus and one that has been associated with FT more than any other is *ya*. Wilson (1999:81, 87) analyzes *ya* as a stative verb (when occurring alone) and as a marker of finiteness (when in combination with other verbs). It is not clear what finiteness is defined as since there is no overt marking of tense, aspect or mood.

Furthermore, *ya* is not a copula and all finite verbs are preceded by *ya*. The stative verb-*ya* can denote location/existence (‘there is/there are’) (Wilson 1999:88-89). Wilson also notes that there are instances where *ya* marks predicative possession. The distribution and potential function of *ya* is discussed in section 5.3.2.

**Yena.** This marker has also had a lot of attention. Wilson (1999:81) describes it as a stative verb (when alone) and a relativizer. Anon (1916:13) notes that it can replace *être* ‘be’, but that it is more common in relative classes that are preceded by *qui/que* in French. Delafosse (1904:265) notes both *ya* and *yena* as particles of the verb, no further details. *Yena* is discussed in section 5.3.3.

**Reduplication.** (Corne 1999:201) writes that reduplication is a frequent strategy for marking intensity or continuity, using (4) as an example. Reduplication is discussed in this thesis in section 5.3.6.

---

8Why this does not then qualify *ya* and/or *yena* as copula-verbs is not clear.
Français Tirailleur (FrTir0328, Anon 1916)

**Original**

*tirailleur* ya besoin tirer, tirer, tirer toujours

**Stdzd**

tirailleur ya besoin tirer, tirer, tirer toujours

**Gloss**

tirailleur ya need shoot shoot shoot always/stretching

‘The tirailleur should shoot without stopping’

**Gagner.** (Delafosse 1904:265) suggests that *gagner* ‘win/acquire/get’ can mean ‘become’ and mark past tense, but it is not clear how strong his claim is. (Wilson 1999:91) notes that *ya gagner/gagné* can form a passive construction. Section 5.3.4. of this thesis contains details on *gagner* in the corpus.

**Content.** (Wilson 1999:53) writes that *ya content* means ‘want to’ (as opposed to French *vouloir* ‘want’), this is dealt with in section 5.3.5.

**Statements not investigated directly in this thesis**

**Verbalizer.** New verbs can be created using the verbalizer, *faire* ‘do’ (Wilson 1999:86 and Corne 1999:200). The construction *faire manière* + a verbal element means ‘to try’. In combination with a nominal element *faire manière* means ‘to use’. If it precedes a clause it means ‘to act in such a way as to’ Wilson (1999:86-87).

The construction *faire mêmechose* is a periphrastic construction that means ‘to act like something’. It is a strategy used to overcome gaps in vocabulary (Wilson 1999:87 and Corne 1999:200).

**Demonstratives.** Demonstratives pronouns are expressed by *ça* ‘that/this’ or .. *yen à là* (roughly ‘that which is there’). Example: *ça tirailleur* or *tirailleur yena là* (Anon 1916:9 and Wilson 1999:34). Delafosse (1904:265) also describes the use of *là* ‘there’ as demonstrative.

**Prepositions.** There are few, if any, prepositions (Anon 1916:14 and Wilson 1999:43). The preposition *à* (and therefore also the portmanteaus *au* and *aux*) is deleted (Delafosse 1904:265). Example: *je vais au village* becomes *moi parti village*.

**Coordination is asyndetic.** Conjunctions are often, if not always, omitted (Wilson 1999:93-94).

**Nominal morphology.** There is no singular/plural distinction on nouns (Anon 1916:8 and Wilson 1999:31).

**Verb morphology.** The main verb is always in what is described as “the simplest form”. This often what is in the lexifier French called the infinitive form (Anon 1916:12 and Wilson 1999:39), but (Delafosse 1904:265) notes that it can also be the past participle or imperative form of French. This means that there is no inflection on the verb for person, gender or number.

**Mood.** Wilson (1999:106) writes that there is no formal distinction of mood. Imperative clauses are however often subject-less (Wilson 1999:91 and Corne 1999:201).
Aspect Wilson (1999:106) writes that there is no marking of aspect. (Delafosse 1904:265) describes a verbal construction that could be interpreted as perfective past or present perfect. It is formed with a/ya gagné + past participle. Example: lui ya gagné mort meaning 'he is dead'.

3.2.4 The linguistic makeup of the Français Tirailleur troops

There are many languages that have been mentioned in connection with FT. First of all there is great consensus on the matter of the lexifier (the language that has contributed with most lexical items), it is French (as spoken in 1800’s). The primary contact with the French language must have been through spoken discourse, which means that it is unlikely that structures and words that were uncommon in the spoken language at that time could have made it into the pidgin.

There was a French officer by the name of Charles Mangin in the late 19th century and beginning of the 20th century. In 1910 he published a book called La Force Noire in which he propagates for the use of African troops in the event of a European war. This book also contains valuable information about the African troops and their composition.

Mangin (1910:274-275) writes that the first troops of African soldiers were mainly composed by the Wolofs and Toucouleurs. The Toucouleurs⁹ are a part of the Fulani group.

⁹The term ‘Toucouleur’ is French popular etymology based on the fact that these people are believed
Both the Wolofs and the Fulani speak languages of the North-Atlantic branch of the Niger-Congo family and are predominately muslim. The Fulani group is a large ethnic group in West Africa that has a wide geographical spread and many different dialects/languages (Harrison 2003). We cannot be sure what variety of the Fula was spoken by these West African soldiers.

Biondi (1987:49) points out that one of the biggest differences between slavery in West Africa and the new world (the americas) was the presence and importance of the “mixed” population and ‘signares’ in particular. ‘Signares’ is a term used for African or part-African women who were companions to the French men of the colony.

Thus, the first troops were made up of speakers of Wolof and Fula. It is not unlikely that there were soldiers, speakers of Fula in particular, who had some knowledge of French prior to enrollment.

After Wolof and Fula, the Serer (also Northern Atlantic; Niger-Congo) were added to the troops. At the time of publication of La Force Noir the most dominating group was Mande (Bambara, Mandinka, Mende, Dyula, Soninke and Susu) and they were recruited after the Serer. The languages spoken by the Mande group are not related to the Niger-Congo. The Bambara and Mandinka were recruited first and later Susu and Dyula\textsuperscript{10} and lastly Soninke.

The last group that Mangin notes that the French military recruited were the Hausa of Dahomey (modern day Northern Benin). The Wolof and Toucoulers (Fula) were preferred in the beginning according to Mangin because they were easier to incorporate in the military as there were already indigenous officers who spoke their languages (Wolof and Fula). The Bambara proved more difficult to instruct at first since they did not speak Wolof nor Fula, but they did later make up the largest group of the West African troops in the French colonial army.

The Hausa language is not related to Niger-Congo or the Mande languages, but a member of the Afro-Asiatic language family. This means that there were languages from at least three separate language families spoken among the soldiers: Niger-Congo (Northern Atlantic), Mande and Afro-Asiatic.

Van Den Avenne (2012:258) underlines the influence of Bambara in FT; she describes the anonymous manual from 1916 as portraying FT as a “calque\textsuperscript{11} de la langue Bambara”.

Table 2 contains information on the different languages that have been mentioned in connection with FT, primarily by Mangin. The genealogy is taken from WALS and the information on geography from Ethnologue. We cannot be sure that all of these languages are indeed involved in FT, but they have been proposed as such.

\textsuperscript{10}Dyula” is not to be confused with Dioula-Fogny, another Niger-Congo language, but of the Northern Atlantic branch.

\textsuperscript{11}\textit{Calque} is a linguistic term, mainly used in French literature. It denotes a word-by-word translation-loan where each lexical element is replaced (in this case by a French word) but the structure is the same. In this context it is possible to say that she describes the FT of the manual as relexified Bambara.

to originate from the state of ‘Takrur/Tekrur/Tekrour’, a West African state in 800 - 1285. The French have reinterpreted that name as meaning ‘Toucouleur’ - all colors.
<table>
<thead>
<tr>
<th>Language Family &amp; Genus</th>
<th>Geographical spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>Indo-European, Romance</td>
</tr>
<tr>
<td></td>
<td>France &amp; West Africa</td>
</tr>
<tr>
<td>Wolof</td>
<td>Niger-Congo, Northern Atlantic</td>
</tr>
<tr>
<td></td>
<td>Senegal, Gambia, Mauritania</td>
</tr>
<tr>
<td>Serer/Noon</td>
<td>Niger-Congo, Northern Atlantic</td>
</tr>
<tr>
<td></td>
<td>Senegal</td>
</tr>
<tr>
<td>Fula (Senegal)</td>
<td>Niger-Congo, Northern Atlantic</td>
</tr>
<tr>
<td></td>
<td>Senegal, Mali, Sierra Leone, Guinea</td>
</tr>
<tr>
<td></td>
<td>Burkina Faso, Mali, Togo</td>
</tr>
<tr>
<td>Môoré</td>
<td>Niger-Congo, Northern Atlantic</td>
</tr>
<tr>
<td>Dyula</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Mali, Côte d’Ivoire, Burkina Faso</td>
</tr>
<tr>
<td>Bambara</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Mali, Côte d’Ivoire</td>
</tr>
<tr>
<td>Mandinka (Gambian)</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Senegal, Gambia, Guinea-Bissau</td>
</tr>
<tr>
<td>Mandinka (Senegal)</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Senegal, Gambia, Guinea-Bissau</td>
</tr>
<tr>
<td>Maninkakan (Western)</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Senegal, Gambia, Malo</td>
</tr>
<tr>
<td>Mende</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Sierra Leone, Liberia</td>
</tr>
<tr>
<td>Susu</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Guinea, Sierra Leone</td>
</tr>
<tr>
<td>Soninke</td>
<td>Niger-Congo, Western Mande</td>
</tr>
<tr>
<td></td>
<td>Mali, Gambia, Guinea-Bissau, Mauritania, Senegal</td>
</tr>
<tr>
<td>Hausa</td>
<td>Afro-Asiatic, Western Chadic</td>
</tr>
<tr>
<td></td>
<td>Nigeria, Benin, Burkina Faso, Cameroon, Ghana, Niger, Sudan</td>
</tr>
</tbody>
</table>

Table 2: Languages proposed as involved in the multilingual situation of FT.

3.2.5 Summary: Les Tirailleurs Sénégalais, West Africa and France

West Africa became known to European powers in the 14th century. France had a great influence over the region during the 17th and 18th hundreds, after the ‘Scramble for Africa’ France was in possession of great territories inland as well.

The French colonial army enrolled native people of Africa, the soldiers of West Africa were called Les Tirailleurs Sénégalais. The West African troops were made up of people of many different ethnic origins, they did not share a common language.

There is mention of a contact language variety derived from French in these troops and by 1916 the French military even publishes a manual of the pidgin. Most previous literature have referred to this manual. It contains much information, both explicitly and implicitly. There is also the work of Maurice Delafosse describing FT.

Delafosse (1904), Wilson (1999), Corne (1999) and the anonymous manual describe a reduced language variety which lacks morphology and differs from the lexifier French in

\[\text{RDC} = \text{République Démocratique du Congo}\]
certain respects. The anonymous manual also notes that there is substrate influence, primarily from Bambara. FT is said to lack prepositions, conjunctions and question particle. There are two forms that are very frequent and characteristics: ya and yena. The literature is not clear on the nature of these forms; they are described as stative verbs, markers of finiteness and relativizers.

FT is said to lack overt copula and have different strategies of filling lexical gaps, such as periphrastic constructions with verbalizers.

There are many West African languages that are likely to have been present in the multilingual situation that gave rise to FT, it would seem, however, that we have reason to believe that Bambara, Fula and Wolof might be particularly influential.
4 Method

This is a corpus study of manually annotated material. The corpus was collected, translated, standardized and glossed by Mikael Parkvall. The annotation has been done by me. Mikael Parkvall is a well-known researcher and has published many articles and books on contact languages, among others his PhD dissertation on substrate influences on Atlantic Creoles (Parkvall 2000).

Each string is accompanied by meta-data: bibliographical information and notes on date of production. If the original source contains any information about the speaker (ethnicity or gender) this has also been included.

One important issue is the spoken nature of FT. Pidgins are less often put into print and many of the utterances found in the corpus are representations of speech as opposed to written text that was intended to be read. In a spoken discourse, it is possible for the speaker to make use of situation specific information, shared knowledge or gestures. All this extra information is lost in writing and this makes the material more difficult to work with.

4.1 Source material

This section contains information about the documents that constitute the corpus, special attention is devoted to three documents that have contributed greatly.

This corpus contains material from 148 different documents by 130 authors. The oldest document was published in 1864 (Dupratz 1864) and the most recent 2008 (Ruault 2008). The date of publication is not necessarily the same date as the author attributes to the production of the utterance. This means that while Dupratz still contains the oldest utterances, but the utterances with the most recent date of production are found in Biasini (1995) and Bellaigue (2009). These are attributed to the 1950’s.

Table 3 shows the strings of the corpus as they are distributed over time of utterance. Several documents contain material that is attributed to several different dates. Table 4 shows the eight documents that have contributed the most to this corpus. The two sections DESC and COMM of the anonymous manual have been treated as two different documents by two different authors.
Words | Strings | Authors | Documents
--- | --- | --- | ---
Total | 8 686 | 1 737 | 130 | 148
Before 1916 | 4 564 | 945 | 82 | 102
After 1916 | 4 099 | 788 | 43 | 43
1860-1869 | 67 | 13 | 12 | 16
1870-1879 | 197 | 46 | 5 | 5
1880-1889 | 564 | 116 | 12 | 17
1890-1899 | 2 176 | 448 | 34 | 40
1900-1909 | 765 | 160 | 17 | 19
1910-1916 | 795 | 160 | 21 | 24
1916-1919 | 2 354 | 428 | 14 | 15
1920-1929 | 791 | 155 | 7 | 7
1930-1939 | 447 | 85 | 9 | 9
1940-1949 | 428 | 105 | 11 | 11
1950-1959 | 79 | 15 | 2 | 2
Date unknown | 23 | 4 | 1 | 1

Table 3: FT material sorted by year of production of utterance

Table 4: The eight largest documents of the corpus

<table>
<thead>
<tr>
<th>Document</th>
<th>Words</th>
<th>Strings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anonymous 1916 (COMM)</td>
<td>1505</td>
<td>249</td>
</tr>
<tr>
<td>Cousturier (1920)</td>
<td>573</td>
<td>119</td>
</tr>
<tr>
<td>Baratier (1912)</td>
<td>327</td>
<td>62</td>
</tr>
<tr>
<td>Diarra (1927)</td>
<td>285</td>
<td>48</td>
</tr>
<tr>
<td>Lhote (1947)</td>
<td>253</td>
<td>43</td>
</tr>
<tr>
<td>Marie-Victoria (1921)</td>
<td>211</td>
<td>48</td>
</tr>
<tr>
<td>Desjardins (1925)</td>
<td>201</td>
<td>38</td>
</tr>
<tr>
<td>Leymaire (1898)</td>
<td>197</td>
<td>34</td>
</tr>
</tbody>
</table>

Many of the documents are anthropological in nature, reports from travels by missionaries or military personnel in West Africa. Material written by authors that have not spent any time in the area or met any West African soldiers has been excluded.

The source with the most utterances is the anonymous manual from 1916 (see section 3.2.2.), followed by *Des inconnus chez moi* — a biographical book by Lucie Cousturier.

Lucie Cousturier was not a part of the military; she was a painter and writer. She encountered West African soldiers during WW1 in the French countryside. They were stationed there before going up to the front. She took it upon herself to organize classes in French for them (Little 2009). It was these encounters that formed the basis of her book, *Des inconnus chez moi* ('The Strangers/Unknowns in My Home'). She later travelled to West Africa and wrote a travelogue called *Mes Inconnus chez eux* ('My Strangers/Unknowns in Their Home'), it does not however contain utterances of FT.
The third largest contributing document to the corpus is Albert Baratier’s *Épopées Africaines* — a collection of reports and stories by a French officer who travelled in the area. The accounts of his military life and travels are accompanied by illustrations of Africans and the local wildlife by illustrator Lucien Pouzargues.

The fourth largest contributing document is a letter that was written by a West African soldier — a Tirailleur Sénégalais — by the name of Baba Diarra. He fought for France in the conquest of Morocco in 1912 and in the first World War. During the first World War he became sergeant. The letter is written in FT and concerns the discrimination of African soldiers in the French colonial army. Diarra writes that the West African soldiers are payed less and treated badly (see example 5).

The letter was published in the first issue of the journal ‘La Race Nègre’ in 1927, which was founded by the old West African soldier Lamine Senghor (Dewitte 1990:128). The letter finishes with the suggestion that if war breaks out again it should be fought by soldiers who have the rights of French citizenship and not natives of the colonies.

(5) Français Tirailleur (FrTir0792, Diarra 1927)

Original Gradé ropéens y a compter nous comme sauvasi, comme plus mauvais chien encore

Stdzd Gradé européen ya compter nous comme sauvage, comme plus

Gloss officer European ya count 1pl.ACC as savage, as more mauvais chien encore

bad dog still/again

‘The European officers saw us as savages, even worse than dogs’
4.2 Reliability of the sources

A problem with this study is the reliability of the sources. Authors can be notoriously inconsistent and potentially untrustworthy due to prejudices or lack of insights into the language. Contact language varieties are often associated with a negative and/or exotic stereotype. This can give rise to authors producing utterances that are suspiciously similar to the lexifier or, the other extreme, made more "pidginy" and exotic.

There are utterances included in the corpus that are potentially untrustworthy, it is impossible to know for certain whether or not a author has misunderstood or consciously manipulated the material. Parkvall has in the collection of the material attempted to exclude utterances that are most likely not authentic FT. Two important reasons for exclusion has been (i) excessive glorification of France and (ii) author hasn’t spent any time in the area.

The analysis of the data in the result section is based on utterances from several different sources (unless otherwise indicated), i.e. I want to avoid features that are present in one or very few documents. Most of the features in the result section have not suffered from this problem, the documents are quite alike with respect to what is investigated in this thesis.

4.3 Comparison with other languages

When investigating the structure of FT it is interesting to compare not only with the lexifier of French but also potential substrate languages and linguistic typology.

In order to make this comparison, I have consulted linguistic reference literature and made great use of the World Atlas of Languages Structures (Dryer and Haspelmath 2011). WALS is a large database of grammatical, lexical and phonological features in the languages of the world. The database was constructed by an author team of 55 scholars. The sample of languages varies with the different features. WALS aims at having a balanced sample of the languages of the world, both genealogically and geographically. In order for the different features to be comparable, there is a special set of 100 and 200 languages that are included in most chapters.

It is important to keep in mind that in many WALS chapters, the languages are only coded with one value even though the language might sport more than one strategy for that specific feature. An illustrative example of this is the word order-chapters (Dryer 2011g), where several languages allow for more than one order of subject, verb and object. The word order that is perceived by researchers of the language as dominant is the value found in WALS, even if other types are possible. This is also true of the chapter on expression of polar interrogatives (Dryer 2011h) and many more.

4.4 Limitations of the study – the nature of corpus

The FT-material has been annotated for highly frequent grammatical phenomena such as predicative possession, subject and position of negation. A full account follows in the next section.
The size of a corpus, the nature of the sample and balance determines what research questions it can be used for. When describing low-frequency phenomena, such as the lexicon, a large corpus is needed. This is a small corpus and it is not balanced with respect to speakers, genres, etc. This means that the study must be limited to describing highly frequent linguistic phenomena such as the structure of simple declarative sentences, the pronoun system etc. We cannot and should not attempt at extracting more from the data set than this.

4.5 Annotation of the data

The data has been glossed, translated and annotated by hand to the best of our abilities. The glossing and translation has been done by Mikael Parkvall. The material has been divided into clauses and annotated for structural features has been made by me in the program Microsoft Excel.

Each sentence appears in the corpus in its original form and in a standardized form. The standardized material was created to facilitate comparison and is primarily concerned with coherent spelling of very similar items (for example: émpé → unpeu), deletion of spaces in constructions/words that are most likely one unit (y a → ya, y en a → yena). There are many cases where different verbal forms in orthographic French are pronounced identically, such as the past participle and infinite form. In the standardized material the infinitive form is favored if there is no difference in pronunciation.

This study is not only limited by the small size of the corpus but also by the abilities of non-native speakers to handle the material. I am not a speaker of FT, not a native speaker of French. This is one more reason why we have focused on features that are subject to less variation and ambiguity.

The corpus has been be annotated for

- independent or subordinate clause
- type of sentence (interrogative, declarative, imperative)
- stative or dynamic verb
- expression of polar interrogation
- standard negation (ne... pas, pas, point, mie, goutte)
- non-standard negation (plus 'no more/longer’, rien 'nothing’, jamais 'never’ and personne 'no-one’)
- expression of grammatical gender
- expression of pronouns
- lexical or pronominal subject
• drop of overt subject
• copula relations (predicative, presentative, locative, existential and equative)
• presence of certain potentially interesting specific items (ya, yena, moyen, manière, etc)
• reduplication
• potential TA-marking
5 Results and analysis

This section contains the results and the analysis. It is further divided into four parts. The first part (5.1) concerns the basic makeup of the data; type-token-ratio, independent/subordinate and declarative/interrogative/imperative. The second part (5.2) is labeled “function seeking form” and contains the findings from the investigation of certain basic functions that we expect the pidgin to express. The third part (5.3) is labeled “form seeking function” and is the mirror image of the previous. Instead of investigating the formal expression of an expected function I have attempted at finding a function of a certain highly frequent forms. The fourth and last section (5.4) contains notes on change over time in the material.

5.1 Data makeup

This section covers type/token-counts and basic categories such as independent clause/subordinate clause and stative/dynamic.

As stated previously, the difference between the original and standardized material is primarily more coherent spelling and the deletion of spaces.

5.1.1 Clause-types and stative/dynamic

This section mainly serves as an introduction to the corpus material, one cannot draw conclusions about the structure of FT directly from these counts.

Table 5 shows the amount of independent and subordinate clauses. There are strings that do not consists of clauses but of unique words, these are included in the study and labeled ‘other’ here.

<table>
<thead>
<tr>
<th>Independent clauses</th>
<th>Subordinate clauses</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 521 (88%)</td>
<td>208 (12%)</td>
<td>8 (0.04%)</td>
</tr>
</tbody>
</table>

Table 5: Independent/subordinate clauses

Table 6 contains counts of declarative, interrogative and imperative clauses among the strings. There are strings in the corpus that cannot be categorized in an obvious way, these have not been forced into a category, but marked as ‘other’.

<table>
<thead>
<tr>
<th>Total</th>
<th>declarative</th>
<th>interrogative</th>
<th>imperative</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 737 (100%)</td>
<td>1 602 (92%)</td>
<td>62 (4%)</td>
<td>83 (5%)</td>
<td>90 (0.05%)</td>
</tr>
</tbody>
</table>

Table 6: Declarative/interrogative/imperative/other
Table 7 displays the distribution of clauses into the types ‘dynamic’ and ‘stative’. The stative clauses are further divided into 5 different copula relations (see section 5.3.8) and 1 group of ‘other statives’. Other very frequent statives are ‘know’, ‘be able’ and ‘like’.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Dynamic</th>
<th>All stative</th>
<th>All COP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 737 (100%)</td>
<td>823 (47%)</td>
<td>908 (52%)</td>
<td>498 (29%)</td>
</tr>
</tbody>
</table>

Table 7: Stative and dynamic clauses

5.1.2 Type-token-ratio in the corpus

A **token** is one occurrence of one word. A **word** is here defined as a number of characters surrounded by spaces\(^{12}\). A **type** is the the total occurrences of a word that is spelled the same \(^{13}\). The data is not lemmatized, types are defined by their orthographical expression.

The original material contains 10 083 tokens and 1 464 types. If we exclude *hapax legomena*\(^{14}\) there are 705 types. The standardized material consist of 8 696 tokens and 1 172 types (629 without hapax). This material has been divided into 1 737 strings. There is material included in the corpus which does not consist of full clauses, there are 62 such strings.

Pidgins are said to be characterized by a restricted vocabulary (David 1971:15 and Holm 1988:73). Anon (1916) even writes that it is important to avoid synonyms when communicating with the West African soldiers.

> “Mais il faut également [...] réduire le plus possible le nombre des mots employés et, par conséquent, éviter d’exprimer la même idée par plusieurs mots différents, ce qui dérouterait l’indigene et lui rendrait la compréhension de notre langue très difficile.” (Anon 1916:17)

Translation into English by Wilson (1999:48)

*But equally necessary are [...] to reduce as much as possible the number of words used and thus avoid using several different words to express the same idea, as this will confuse the native and make it hard for him to understand our most difficult language.*

If speakers of FT use less synonyms and have a restricted lexicon we would expect this to show in the corpus material. One way to measure the lexical diversity in a corpus is through the type-token-ratio (TTR). TTR is the number of types divided by the number of tokens. A low TTR would suggest that a few types are being used very often. This could

\(^{12}\)The amount of tokens will differ between the original material and the standardized since part of the standardization involved the deletion of spaces.

\(^{13}\)This means that *she loves you, yeah, yeah, yeah* for example contains 6 tokens but only 4 types.

\(^{14}\)Hapax Legomena are types with only one occurrence.
be due to a small corpus, limited lexicon or little/no morphology\textsuperscript{15}. This measurement tends to be even lower for spoken corpora.

In order to compare small and large corpora fairly one can also use a measurement of standardized TTR (sTTR). sTTR is calculated in the same way as TTR, but of 2 000\textsuperscript{16} concatenated words of the corpus (Baker 2006:52).

Table 8 displays the amount of types, tokens and type-token-ratios of the FT-corpus. The three most frequent types have also been included. ‘Type (≥2)’ refers to the number of types that have 2 or more occurrences, i.e. all types minus *hapax legomena*.

\begin{table}[h]
\centering
\begin{tabular}{lllllll}
\hline
Token & Type & Type (≥2) & TTR & sTTR & Freq 1 & Freq 2 & Freq 3 \\
\hline
Original & 10 082 & 1 464 & 705 & 14\% & 27\% & a (866) & y (674) & moi (355) \\
Stdzd & 8 696 & 1 172 & 629 & 13 \% & 26 \% & ya (682) & moi (357) & pas (318) \\
Gloss & 8 765 & 1 052 & 666 & N/A & N/A & sg (940) & ya (670) & NEG (339) \\
\hline
\end{tabular}
\caption{Types/tokens in FT}
\end{table}

A sTTR of 27\% is quite low, but perhaps not as low as one might have expected. We cannot compare sTTR directly across languages, but here is brief comparison with English to give the reader some point of reference. English is not an isolating language (Bickel and Nichols 2011) and the sTTR of of the spoken material of the British National Corpus is 33\% Baker (2006:52). We need more data to compare corpus’s properly, and ideally we would like to compare FT with spoken French of the 19th century.

FT, like many (if not most) pidgins, seem to lack morphology, which would pull down the sTTR. The material in this corpus is most likely primarily spoken and it is small. All of these factors should indicate a low sTTR, but perhaps 27\% isn’t as low as we expected. This is however very speculative and we need more detailed data on the distribution of tokens and types in spoken informal French and other pidgins to further this issue.

Another way of investigation types and tokens in corpora is by the application of Zipf’s Law. Zipf (1935) found that many natural phenomena follow a pattern where the second most frequent item is half as frequent as the most frequent, the third half as frequent as the second most frequent, and so on.

This formula is known as *Zipf’s Law* and it is predicted to be true of such different things as size of population in cities and word frequencies in language. In language, it seems to work quite well (often with the exception that the first 100 items are often irregular).

Many corpora of natural language follow this pattern, but does it also apply to pidgins in general and FT in particular?

Figure 5, 6 and 7 illustrate the type frequency of the material\textsuperscript{17}. The x-axis is the types, and the y-axis the frequency of those types. The types are sorted after frequency on the

\textsuperscript{15}Provided we are defining types by their orthographical/formal expression and not by their lemma-form.

\textsuperscript{16}Some authors use 1 000, the default value of the latest version of WordSmith. 2 000 is used in this study.

\textsuperscript{17}Keep in mind that there are fewer types in the standardized material as opposed to the original.
The figures display the original source material (red), the standardized material (blue) as well as what the perfect Zipf’s curve would be (black).

Figure 5 displays the frequency of all types, figure 6 of only the first 100 and figure 7 is all types logarithmically. All frequencies are normalized and both standardized and original material are significantly similar to the perfect Zipfs (p<0.05).

Figure 5: Type-frequency-ratio of the FT-corpus. X-axis = types ranked by freq, Y-axis=frequency. Non-Logarithmic.
Figure 6: Type-frequency-ratio of the FT-corpus. Non-logarithmic, zoomed in at first 100 types.

Figure 7: Type-frequency of the FT-corpus. X-axis = types ranked by freq, Y-axis=frequency (log-log=10)
5.1.3 Summary: data makeup

There are more independent clauses than subordinate, more declarative than interrogative or imperative. We cannot derive structure of FT directly from these numbers, but we can deduce that we have more data for describing declarative independent clauses than any other type of clause.

An analysis of TTR of the material reveals that further investigation into the lexicon of pidgins is needed. Despite the anonymous manuals advice to avoid synonyms and the overall characteristics of pidgins to have a restricted lexicon, the TTR is perhaps not as low as could have been expected.

Zipf’s Law applies to this corpus, even though it is a small corpus of a reduced variety.

5.2 Function seeking form

5.2.1 Polar Interrogative

There are two chapters in WALS (92 & 116) devoted to the expression of polar interrogatives, both written by Matthew Dryer. Chapter 116 concerns the expression of polar interrogatives and 92 covers the position of the question particle (if there is one).

The most common strategy found in these languages is the question particle (61%, 584/954 languages). 130 of these languages have the question particle in initial position and 313 in final. The second most common strategy is interrogative intonation only (18%, 174/954). In Dryer’s study, a language is only coded as having ‘intonation only’ if that is indeed the only means of expressing polar interrogation. This suggest that many (if not most) languages make use of applying question intonation\(^{18}\) to declaratives in order to form a polar interrogative even if they have other strategies as well. Dryer notes one exception to this, there is not interrogative intonation in Imbabura Quechua. He also comments that it is very likely that the type ‘intonation only’ is underrepresented in his survey.

The third most common strategy is interrogative verbal morphology (17%, 164/954). There is only one language in the survey that reportedly has no distinction between declaratives and interrogatives, and this is Mixtec (Chalcatongo) of the Oto-Manguean family.

It is interesting to note that the strategy used by several Germanic languages, but also Spanish and French, — change of word order (7) — is typologically quite uncommon. Dryer found that 13 languages (1%) of his balanced sample of 954 used this strategy.

French French has three main strategies for expressing polar interrogatives.

\(^{18}\)Interrogative intonation does not necessarily mean rising intonation.
French, interrogative intonation (own example)

tu veux manger ça?
2sg want.PRS eat.INF DEF

‘Do you want to eat that?’

French, change of word order (own example)

veux tu manger ça?
want.PRS 2sg eat.INF DEF

‘Do you want to eat that?’

French, initial question particle: (own example)

Est ce que tu veux manger ça?
COP INDEF REL 2sg want.PRS eat.INF DEF

‘Do you want to eat that?’[Lit: ‘Is it so that you want to eat that?’]

Dryer (2011h) has coded French as having a sentence initial particle since the interrogative intonation is not the only strategy and the change of word order is very infrequent (Harris 1988).

Several of the languages the might be involved in the origin of FT are found in Dryer’s chapters on polar interrogatives in WALS. Table 9 presents an overview of the expression of polar interrogation in certain West African that potentially involved in the multilingual situation of FT. All data is from Dryer (2011h) and Dryer (2011j) unless otherwise indicated.

<table>
<thead>
<tr>
<th>intonation only</th>
<th>Baoule, Diola-Fogny, Koyra Chiini (Songhay), Lingala and Mende</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial question particle</td>
<td>Wolof</td>
</tr>
<tr>
<td>final question particle</td>
<td>Bambara, Fula (Cameroon), Hausa, Maninkakan (Western), Môoré, Serer/Noon, Soninke and Fongbe (Lefebvre and Brousseau 2002)</td>
</tr>
<tr>
<td>question particle in either position</td>
<td>Koyraboro Senni (Songhay) and Malagasy</td>
</tr>
</tbody>
</table>

Table 9: Polar interrogatives in certain West African languages

Most of the West African languages in table 9 employ a particle, primarily in a final position.

Out of the 1737 clauses in the corpus of FT, 43 are polar interrogatives. 37 of those are most likely examples of ‘intonation only’. They appear to be identical to the declarative sentences with the exception of a question mark — leading us to believe that there is interrogative intonation. As we have no audio recordings of FT we cannot be certain. There is no significant change over time nor is there any significant variation across authors.
There is however also the issue of subject drop (9), 18% of the FT clauses lack overt subject marking. This means since change of word order in French involves the overt expression of the subject, it is impossible to distinguish between ‘intonation only’ and ‘change of word order’. There are 7 instances of subject drop in the polar interrogatives.

(9) Français Tirailleur (FrTir0051, Cousturier 1920:35)

Original Y a pas content avec nous?

Stdzd Ya pas content avec nous?

Gloss ya NEG be.happy/like with 1pl

‘Don’t you like us?’

While there is no way of knowing where this dropped subject is positioned I believe that the most appropriate way to treat these cases are that they display interrogative intonation. There is a lot of subject drop in non-interrogative clauses as well, making the only formal distinction the presence of a question mark.

There are two clear examples of word order change in expressing polar interrogatives. The first is said to have been uttered by a native speaker of Wolof (10) and the second by a white native speaker of French (11).

(10) Français Tirailleur (FrTir0845, Sonolet 1911:44)

Original Dis, mon colonel, veux-tu du cadeau de Patey-Sar?

Stdzd Dire, mon colonel, veux-tu du cadeau de Patey-Sar?

Gloss say my colonel want-you ART gift of Patey-Sar?

‘Say, colonel, would you like a present from me?’

(11) Français Tirailleur (FrTir0041, Cousturier 1920:40)

Original Y a toi content pour lancer grenade?

Stdzd Ya toi content pour lancer grenade?

gloss ya 2sg.NOM be.happy/want for throw grenade

‘Do you enjoy throwing grenades?’

These two instances of word-order change are most likely very rare ways of marking polar interrogation in FT and the fact that one was uttered by a native speaker of French makes it very suspicious indeed.

These is no example of the lexifier particle est-ce que. Question particles are typologically common and present in most of the potential substrate languages, including presumably much influential Wolof, Bambara, Fula and Maninkan. However, the question particle of French is sentence initial. Only three of the potential substrate languages of FT in Dryer’s sample employed initial particle.
It would seem that FT employs ‘intonation only’ to express polar interrogation, but we cannot be certain. Anon (1916:16) and Wilson (1999:47) writes that FT uses ‘intonation only’, but in order to be certain we need to consult recordings (most likely impossible).

We do not know how many language employ intonation when expressing polar interrogation, we only know how many use intonation as the only strategy. Perhaps further light can be shed on the use of intonation in pidgins when The Atlas of Pidgin and creole Language Structures is released, it will feature a chapter on the expression of polar interrogation.

5.2.2 Standard Negation

The most common strategy for expressing Standard Negation in (SN) in FT is with the pre-verbal particle *pas* (12). Standard negation is the negation of an entire verbal declarative main clause (Miestamo 2007).

(12) François Tirailleur (FrTir1182, Barret 1888:324)

<table>
<thead>
<tr>
<th>Original</th>
<th>Stdzd</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moi pas mire lui</td>
<td>Moi pas mire lui</td>
<td>Moi 1sg.NOM NEG mire 3sg.ACC</td>
</tr>
</tbody>
</table>

‘I can’t see him’

If we look at the languages of the world, we find that particles are the most common strategy. In the typological survey by Dryer (2011c) 43% (504/1159 languages) employed a particle. The second most common strategy is a negative affix on the verb, 34% (396/1159), and 10% (120/1159) had some type of Double Negation (DN). The most common position of the negative element in this survey of the world’s languages is pre-verbal, 40% (524/1326) (Dryer 2011e).

French has two strategies for expressing SN, double negation construction of *ne... pas* and post-verbal *pas* alone. This double negation construction is older, *ne* is the former standard negator and *pas* is etymologically a word meaning ‘step’. *Pas* was at first used to reinforce the negation as a minimizer (‘I won’t walk, not a step’). In the 16th century it was also possible to use *point* (‘dot’), *mie* (‘crumb’) and *goutte* (‘drop’) in this construction, but *pas* was most frequent.

In modern spoken French, it is possible to express SN with post-verbal *pas* only (Hopper 1991:26-27), this diachronic change is known as ‘Meillet’s Spiral’ or ‘Jespersen’s Cycle’.

The particle *pas* is also the negator of non-verbal clauses such as possession and copula relations. The particle occurs after the finite verb, but before the main verb and before the object (if it is not a pronominal object). The typical order can be described as:

subject-(ne)-(pronominal object)-finite verb-pas-(main verb)-(nominal object).

If there are no auxiliaries the main verb is the finite verb.
The order of the optional double negation construction in French (S(Neg)VNegO) is not the most common order of double negation in SVO languages (Dryer 2011d). It is more common that the second negator follows the object (SNegVONeg).

If we look at some of the West African languages that are potentially interesting for the development of FT in table 10, a majority of them employ a particle.

<table>
<thead>
<tr>
<th>Pre-verbal Negative particle</th>
<th>Mende, Hausa, Soninke, Koyraboro Senni (Songhay), Malagasy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-verbal Negative particle</td>
<td>Lingala, Gba ya Kara</td>
</tr>
<tr>
<td>Negative affix</td>
<td>Diola-Fogny, Fula (Cameroon), Serer/Noon, Fula (Nigeria)</td>
</tr>
<tr>
<td>Pre-verbal Negative word, unclear if verb or particle</td>
<td>Koyra Chiini (Songhay), Wolof, Fongbe</td>
</tr>
<tr>
<td>Obligatory Double Negation</td>
<td>Baoulé, Môoré, Kikongo</td>
</tr>
</tbody>
</table>

Table 10: Expression of Standard Negation based on Dryer (2011c,i,e).

Wolof has more than one construction, a pre-verbal negative word and a negative suffix. The obligatory double negation of Baoulé is, interestingly, formed with negative tone and a post-verbal negative particle.

The most common way of expressing SN in the FT clauses is by means of the particle *pas* before the predicate. If the clause contained *ya, yena, avoir* or *être* the negator precedes these. There are also a few instances of the French double negation construction *ne... pas*. Table 11 displays the expression of SN in FT.

<table>
<thead>
<tr>
<th>Negator</th>
<th>Total freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total SN</td>
<td>324</td>
</tr>
<tr>
<td><em>pas only</em></td>
<td>297</td>
</tr>
<tr>
<td><em>n’est pas</em></td>
<td>1</td>
</tr>
<tr>
<td><em>n’étais pas</em></td>
<td>1</td>
</tr>
<tr>
<td><em>n’a pas</em></td>
<td>12</td>
</tr>
<tr>
<td><em>n’y a pas</em></td>
<td>9</td>
</tr>
<tr>
<td><em>n’y en a pas</em></td>
<td>2</td>
</tr>
<tr>
<td><em>point</em></td>
<td>0</td>
</tr>
<tr>
<td><em>goutte</em></td>
<td>0</td>
</tr>
<tr>
<td><em>mie</em></td>
<td>0</td>
</tr>
</tbody>
</table>

Table 11: Expression of Standard Negation in FT

There is no evidence of *point, mie or goutte* — *pas* dominates totally. *Pas* occurs as the sole negator in many different sources and from very early on. This is expected, the
other negators were already very infrequent in the lexifier. There is no doubt that *pas* is the most frequent marker of SN.

With one exception, all of the instances of *pas* are directly before the predicate (13) but after the stative verbs/auxiliaries/predicate markers of *ya/yena/avait/étais* (if they are present). This is different from the lexifier, where *pas* occurs after the finite verb, even if it is the only verb in the sentence (14).

(13) François Tirailleur (FrTir0182, de Mandat-Grancey 1900:134)

*Original* Toi dire toujours les noirs *pas* pouvoir suivre horaire!

*Stdzd* Toi dire toujours les noir *pas* pouvoir suivre horaire!

*Gloss* 2sg.NOM say always/still ART black NEG be.able follow timetable

‘You always say that black people cannot follow a timetable!’

(14) French (own example)

Tu dis toujours que les noir-s (ne) peuvent *pas* suivre les

2sg say always that ART black-PL (NEG1) be.able NEG2 follow ART

horarie-s
timetable-PL

‘You always say that black people cannot follow a timetable!’

There is one exception where *pas* occurs after the verb (15), but it is only one out of 297 instances of *pas*.

(15) François Tirailleur (FrTir0308b, Bouquet and Hosten 1917)

*Original* ti manges pas charognards

*Stdzd* tu mange pas charognard

*Gloss* 2sg.NOM eat NEG vultures

‘You don’t eat vultures’

The authors of the original documents have represented the following expressions with accents and spaces, suggesting that they are compositional: *n’a pas, n’y a pas* and *n’y en a pas*. However, if we consider the translations it is more likely that they could be analyzed as entire words, consider (16).

36
If *n’a pas* did indeed consist of DN and the verb *avoir* it is less likely that the translation would be in the present tense. This is the case for most of the 25 occurrences of *n’a pas*, *n’ y a pas* and *n’ y en a pas* which suggests that these expressions might be three distinct particles in their own right. There are however not many instances of this, the most common expression of SN is still “*pas only*’ and these alternative constructions occur mostly in the older material.

Pre-verbal negation with *pas only* is the overwhelmingly most frequent strategy for expression Standard Negation in FT. There are other minor strategies, but they are not nearly as frequent as pre-verbal *pas*.

The findings are in accordance with previous statements by Wilson (1999), but contradict Delafosse (1904:265) who claim that *pas* is post-verbal.

### 5.2.3 Non-standard negation (*jamais, rien, personne, plus*)

Non-standard negation is here defined negatively, i.e. by not being SN (see section 5.2.2). SN might in some languages be used to express the negation of possession, time, existence etc, but it is common than languages have special strategies for this (see Veselinova 2010 and Veselinova 2013). In the case of French and FT, the relevant functions are the negation of time (*jamais– ‘never’ and *plus – ‘no longer’); and negative indefinite pronouns (*rien, personne – ‘nothing’, ‘no-one’*).

Negation of non-verbal clauses (copula, possession, existence, location) is expressed with the same means as SN, pre-verbal *pas*.

In written French, this type of negation is expressed with a double negation construction involving the old negator; *ne* before the finite verb and *jamais, plus, rien* or *personne* after: *je ne regrette rien* (‘I don’t regret anything’). *Jamais* – ‘never’ and *plus* – ‘no longer’ are defined as adverbs, *rien* – ‘nothing’ and *personne* – ‘no-one’ as negative indefinite pronouns. It is not ungrammatical to omit *ne* in spoken discourse (Alsenoy 2011).

As negative indefinite pronouns *personne* and *rien* can also occur in sentence-initially as subjects with *ne* directly after; *personne ne me connaît ici* (‘No-one knows me here’) and *rien ne change* - ‘nothing changes’.

Table 12 displays the occurrences of non-standard negation in FT.
Most of the instances of these words in the FT material are very similar (if not identical) in function to the lexifier. It is however very common to omit *ne*, i.e. *rien, plus, jamais* and *personne* are inherently negative. As with the expression of SN, the instances of the double-word-construction with *n'/ne* are scarce and all before 1900.

Omission of *ne* is possible in modern spoken French in many of these instances. The omission of *ne* is most likely a reflection of the spoken French at that time rather than a process that took place in both FT and French independently.

Both *rien* and *personne* have non-negative function in French. *Rien* is etymologically a positive indefinite noun meaning ‘thing’ and there is still a marginal use of *rien* meaning ’very small thing’ (cf. Eng. ’ounce’) in modern French (‘il restera souvent un rien de suspicion - ‘there remains an ounce of doubt still’). *Personne* can also mean ’person’ in modern French.

There are also occurrences of *plus* in the non-negative meaning of ’more’ in the FT-material, this could potentially be a problematic ambiguity.

The expression of non-standard negation in FT is very similar, if not identical to the expression in the lexifier French. There are however not many examples of non-standard negation, perhaps they differ more than this material suggests.

### 5.2.4 Grammatical gender

Some languages divide nouns into different groups called ‘grammatical gender’ or ‘noun classes’. The term ‘noun class’ is sometimes used when there are more than two or three of these groups. I will use ’grammatical gender’ and ’noun class’ synonymously.

The most important feature of grammatical gender is agreement; elements such as articles, demonstratives, numbers have to agree with the gender of the noun they modify (Corbett 2011).

French divides nouns into either masculine or feminine. The masculine gender in French is overall more frequent and inanimate nouns are mostly masculine. French has

---

Table 12: Expression of non-standard Negation in FT

<table>
<thead>
<tr>
<th>English equivalent</th>
<th>Negator</th>
<th>Total freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>no-one</td>
<td>personne</td>
<td>5</td>
</tr>
<tr>
<td>never</td>
<td>jamais</td>
<td>19</td>
</tr>
<tr>
<td>never</td>
<td>*n ya /... */ jamais</td>
<td>3</td>
</tr>
<tr>
<td>nothing</td>
<td>rien</td>
<td>17</td>
</tr>
<tr>
<td>nothing</td>
<td>*pas /... */ rien</td>
<td>2</td>
</tr>
<tr>
<td>no more</td>
<td>plus</td>
<td>10</td>
</tr>
<tr>
<td>no more</td>
<td>*n ya /... */ plus</td>
<td>2</td>
</tr>
<tr>
<td>no more</td>
<td>n’a plus</td>
<td>1</td>
</tr>
</tbody>
</table>

---

19From the Latin accusative of *res* meaning ’thing’
20The relation between the “natural” gender of masculine and feminine and these grammatical categories is rather bleached. For example: bras are masculine (un soutien-gorge - ‘a bra’).
gender agreement on articles, possessive pronouns, adjectives, adverbs and past participles.

Many languages of West Africa have gender systems consisting of more than two groups (Corbett 2011). Diola-Fogny and Fula (Guinean) are coded in WALS as having ‘five or more’, Hausa has ‘two’ and Bambara none.

Considering that there is little (or no) inflection of adjectives, adverbs or past participles in FT, the place to look for gender agreement is in the few instances of indefinite and definite articles (un/une, le/la) and possessive pronouns (mon/ma, ton/ta, son/sa). There is a total of 322 occurrences of these morphemes in our 1737 strings and they occur all throughout 1860-1950.

Some authors have represented the articles as fused with the noun, such as FrTir0739 émpé - ‘a little’ (Diarra 1927:54). These instances are listed as ‘agglutinated’ in table 13. They are not excluded from the other counts of the articles and pronouns. Cases where the lexifier French agglutinates the article are not counted as agglutination here, such as l’école. Most (80%) of the explicitly agglutinated articles occur in Anon (1916) COMM.

In table 13, “right” and “wrong” refers to whether or not the article or pronoun agrees in gender with the gender of that noun in French. There is no reason to assume the FT has the same genders as French, this is primarily etymological information.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Masculine</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Right”</td>
<td>“Wrong”</td>
<td>“Right”</td>
</tr>
<tr>
<td>Total</td>
<td>322 (100%)</td>
<td>194 (60%)</td>
<td>32 (10%)</td>
</tr>
<tr>
<td>Indefinite articles (un, une)</td>
<td>47 (100%)</td>
<td>11 (23%)</td>
<td>2 (4%)</td>
</tr>
<tr>
<td>Definite articles (le, la)</td>
<td>137 (100%)</td>
<td>58 (42%)</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>Possessive pronouns (mon, ton, son, ma, ta, sa)</td>
<td>128 (100%)</td>
<td>85 (66%)</td>
<td>28 (22%)</td>
</tr>
<tr>
<td>Explicitly agglutinated</td>
<td>44 (100%)</td>
<td>29 (66%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

Table 13: Gender marking in FT

There are more occurrences of words that are etymologically masculine than feminine. More words are assigned masculine gender agreement when they are etymologically feminine than the opposite.

All of the instances where feminine gender has been assigned to a noun that has masculine gender in French (“wrong agreement”) were either ma capitaine (‘my captain’) or ma lieutenant (‘my lieutenant’). This is quite peculiar, especially if we consider that even if there were female officers (of which there is no evidence) they are to be called ‘capitain’ only - the possessive pronoun is only used for male officers (Bobe 2007). These examples may be few, but they are found in five different sources. There sources are not believed to have influenced each other and were all published before 1900.

Anon (1916:7) and Delafosse (1904:265) suggest that the articles can be so agglutinated so that constructions like mon latète are possible. There are four instances of this in
the corpus: FrTir0730 *mon la poule* (Fargeas 1899:131), FrTir0227c *son lamain* (Anon 1916), FrTir0852 *cette la guerre* Costurier (1920:173) and FrTir0366 *ton latète* (Sembène 1987:2h04). 3 out of 7 occurrences of *tête* do not have *la*, 8 of 20 for *guerre* and 9 of 15 for *main*. There is only one instance of *poule*.

There are seven cases of the portmanteau *du (de + le)* and 1 of *au à + le*, all with correct agreement according to French.

As for the statement by Anon (1916) that feminine natural gender on animate nouns is expressed by adding *femme* as a suffix, there is only on sentence displaying this, (17).

(17) Français Tirailleur (FrTir0550, Augouard 1905:231)

**Original** Moi voulait cochons les hommes, pasque moi y en a cochons les femmes, et cochons les femmes tout seuls, y en a pas connaître faire pitit!

**Stdzd** Moi voulait cochon les homme, parceque moi yena

**Gloss** 1sg,NOM want pig ART man because 1sg,NOM yena pig cochon les femme, et cochon les femme tout seul, yena pas ART woman and pig ART woman all aloneyena NEG know make connaître faire petit!

babses

‘I want a boar, because I only have sows, and sows alone cannot have piglets’

However, this seems to be the only instance where the natural gender of an animate being is a relevant topic of the material, so we cannot assume that it is productive, nor that it is non-productive.

Anon (1916:7) claims that all inanimate nouns are treated as having masculine gender. While it is true that masculine gender is more often applied to feminine nouns than the opposite; there are several instances of inanimate nouns with feminine gender marking; FrTir0004 *la bande* - ‘the cartridges’ (Conombo 1989:132).

Out of 1 737 strings there are only 303 forms that mark what is etymologically grammatical gender, and many are given masculine gender though the noun is etymologically feminine. This suggests that FT does not have a productive gender system. This is also proposed in the anonymous manual and explicitly stated by Wilson (1999:110). The majority of the contact varieties with French as a lexifier in APiCS\(^{21}\) do not show gender agreement on adjectives or indefinite articles (Maurer and the APiCS Consortium 2013).

### 5.2.5 Personal pronouns

The French personal pronoun system is rather complicated and makes distinctions between nominative, accusative and dative case. There are also reflexive forms and so called “dis-
junctive” pronouns. The disjunctive pronouns are used in prepositional phrases and certain empathic contexts.

Table 14 gives a basic overview of the French pronominal system.

<table>
<thead>
<tr>
<th></th>
<th>NOM</th>
<th>ACC</th>
<th>REFL</th>
<th>DAT</th>
<th>DISJUNCTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>je</td>
<td>me</td>
<td>me</td>
<td>me</td>
<td>moi</td>
</tr>
<tr>
<td>2sg</td>
<td>tu</td>
<td>te</td>
<td>te</td>
<td>te</td>
<td>toi</td>
</tr>
<tr>
<td>3sg</td>
<td>il, elle, on, ça, ce</td>
<td>le, la, en</td>
<td>se</td>
<td>lui, y</td>
<td>lui, elle, soi</td>
</tr>
<tr>
<td>1pl</td>
<td>nous</td>
<td>nous</td>
<td>nous</td>
<td>nous</td>
<td>nous</td>
</tr>
<tr>
<td>2pl</td>
<td>vous</td>
<td>vous</td>
<td>vous</td>
<td>vous</td>
<td></td>
</tr>
<tr>
<td>3pl</td>
<td>ils, elles, les</td>
<td>en</td>
<td>se</td>
<td>leur, y</td>
<td>eux, elles</td>
</tr>
</tbody>
</table>

Table 14: French pronouns

Anon (1916:11), Wilson (1999:37) both claim that the personal pronoun system of FT makes no distinction between nominative case and non-nominative case and indeed this seems to hold up also for material other than the anonymous manual. Table 15 displays the expression of pronouns in FT. There is variation in the material, the table displays the most frequent form followed by absolute number of occurrences and percentage of all occurrences of that function. For example: *moi* occurs 303 times as 1st person singular in subject position (NOM), that is 95% of the times there was a form marking 1st person singular pronoun in subject position in the entire material. The marking of pronouns in what is etymologically accusative and dative case have been lumped together as there was no difference in FT.

<table>
<thead>
<tr>
<th></th>
<th>NOM</th>
<th>ACC &amp; DAT</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td><em>moi</em> 303 (95%)</td>
<td><em>toi</em> 184 (89%)</td>
<td><em>lui</em> 176 (87%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2sg</td>
<td><em>moi</em> 41 (89%)</td>
<td><em>toi</em> 44 (100%)</td>
<td><em>lui</em> 43 (91%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg</td>
<td><em>nous</em> 73 (99%)</td>
<td><em>vous</em> 9 (100%)</td>
<td><em>eux</em> 18 (45%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>nous</em> 21 (100%)</td>
<td><em>vous</em> 1 (100%)</td>
<td><em>eux</em> 2 (100%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 15: Pronouns in FT

As can be seen in table 15 there is typically no distinction between subject and object-forms of the pronouns. The set of pronouns proposed by the anonymous manual does indeed appear in the other documents as well.

There are only 4 instances of reflexive forms, 2 *m’s, 1 nous* and 1 *mon corps* (18).

---

22Including 9 occurrences of *li* as 3sg,NOM
(18) François Tirailleur (FrTir0774, Cousturier 1920:158)

Original Je vas tuer mon corps

Stdzd Je va tuer mon corps

Gloss 1sg.NOM FUT kill 1sg.POSS body

‘I’m going to kill myself’

(Holm 1988:204) discusses instances of words for ‘body’ together with a possessive form meaning ‘self’ in Atlantic creoles. There is evidence of this form in Old French (Sylvain 1936:65 as cited in Holm 1988) and Chaudenson (1974:734 as cited in Holm 1988) has documented its survival in some regional dialects. In a typological study of intensifiers, such as himself/herself, König and Gast (2006:224) found that the expression of these were often derived from words denoting body parts. It is thusly not unsurprising that there is an instance of this in FT, there is however only one.

The distinction in natural gender for 3rd person singular seems to be lost (i.e. elle ‘she’ - il ‘he’). This has been noted for French creoles as well (Holm 1988:201). There are however quite few women mentioned in the material so we cannot be certain that this is really the case. There are three instances of overt pronouns denoting women, all are found in Lhote (1947) (19-20).

(19) François Tirailleur (FrTir0697, Lhote 1947:304)

Original Lui y en a faire cabinet partout dans le campement

Stdzd Lui yena faire cabinet partout dans le campement

Gloss 3sg.NOM yena make toilet all.over in/at the camp

‘She relieved herself all over the camp’

(20) François Tirailleur (FrTir0697, Lhote 1947:305-6)

Original Alors, moi y en a dire Lieutenant et lieutenant y en a boucler lui, car lui y en a cochon beaucoup

Stdzd Alors, moi yena dire Lieutenant et lieutenant yena boucler lui, car lui yena cochon beaucoup

Gloss then 1sg.nom yena say/tell lieutenant and lieutenant yena shackle lui, car lui yena cochon beaucoup

3sg.ACC, because 3sg.NOM yena pig much

‘So I told the lieutenant, and he shackled her, since she is a real pig/asshole’

The pronoun system of FT does indeed appear to be as the anonymous manual claim. No distinction between subject and object position and no natural gender.
5.2.6 Attributive possession

Attributive possession refers to constructions in which the expression of the possessum and possessor form a noun phrase (McGregor 2009:2), such as ‘Kolbeinn’s fox’.

In French, attributive possession where the possessor is pronominal (‘my fox’) is expressed primarily through possessive pronouns\textsuperscript{24} (mon renard). These pronouns agree with the possessum in number and gender and with the possessor in person. Table 16 gives an overview of possessive pronouns in French.

<table>
<thead>
<tr>
<th>Possessor ↓</th>
<th>Possessum →</th>
<th>masculine singular</th>
<th>feminine singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>mon</td>
<td>ma</td>
<td>mes</td>
<td></td>
</tr>
<tr>
<td>2sg</td>
<td>ton</td>
<td>ta</td>
<td>tes</td>
<td></td>
</tr>
<tr>
<td>3sg</td>
<td>son</td>
<td>sa</td>
<td>ses</td>
<td></td>
</tr>
<tr>
<td>1pl</td>
<td>notre</td>
<td>notre</td>
<td>nos</td>
<td></td>
</tr>
<tr>
<td>2pl</td>
<td>votre</td>
<td>votre</td>
<td>vos</td>
<td></td>
</tr>
<tr>
<td>3pl</td>
<td>leur</td>
<td>leur</td>
<td>leurs</td>
<td></td>
</tr>
</tbody>
</table>

Table 16: French possessive pronouns

When there is a nominal possessor in French this is expressed through a prepositional construction (le renard de Kolbeinn ‘Kolbeinn’s fox’). There is also a prepositional construction of à. This construction can be combined with the possessive pronouns giving a slightly different meaning; mon renard à moi - ‘my own fox’.

The anonymous manual describes three different strategies for attributive possession: (i) the possessive pronouns mon and ton for 1st and 2nd person singular respectively, (ii) a prepositional construction with pour and (iii) juxtaposition with possessum preceding the possessor (own example: renard Kolbeinn ‘Kolbeinn’s fox’).

All these constructions exist in the material, along with more prepositional constructions. Table 17 displays the different forms of expressing attributive possession of 1st, 2nd and 3rd singular person and nominal possessors. The most common form of the possessive pronouns is the etymologically masculine form (mon and ton), more on grammatical gender can be found in section 5.2.4.

Plural possessive pronouns are excluded here, there were only two instances of these, notre - ‘our’ and juxtaposition with nous.

\textsuperscript{24}In French grammar tradition they are known as adjectives, but I will label them pronouns here to avoid confusion when discussing the part-of-speech adjectives later.
Table 17: Expression of attributive possession in FT

There is one example of double marking of attributive possession, both pronominal and with the prepositional à (21). It is not clear if this should be interpreted as the typical possessive or the more emphasized ‘own NP’.

(21) Français Tirailleur (FrTir0491a, Hubin 1987
   Original moi y en a ton Noir à toi
   Stdzd moi yena ton Noir à toi
   Gloss 1sg.NOM yena 2sg.POSS black at/to 2sg.ACC
   ‘I am your Black’

It would seem that (Anon 1916) was indeed right that the most common way for 1st and 2nd person singular is possessive pronouns, but also for 3rd person singular. When it comes to nominal possessors it would seem that pour... is less common than expected; in fact the construction with à/aux is more common than de, though they are both very infrequent.

Juxtaposition (possessum - possessor) is the most common strategy in FT when the possessor is nominal. The construction is found in many different sources, the earliest being from 1888 (Binger 1892:207). Holm (1988) notes that juxtaposition is a widespread strategy in creoles, though primarily with the order possessor - possessum. Holm also suggests that the juxtaposition construction with the order possessum - possessor in French creoles might be due to the deletion of prepositional de (renard de Kolbeinn → renard Kolbeinn). Whether or not this is the case for FT is not known, but likely.

Possessum - possessor is also the most common order among the French lexified contact varieties in APiCS, and it is slightly more common than ‘possessor - possessum’ in all the contact varieties in APiCS (Huber and the APiCS Consortium 2013).

It would seem that attributive possession where the possessor is pronominal is most often expressed through possessive pronouns. Juxtaposition and the prepositional construction with de are equally common when the possessor is nominal. The order in the juxtaposition is interestingly not as the anonymous manual suggests (possessor - possessum), it would appear that the substrate influence is perhaps less than the author(s) of the manual thought.
5.2.7 Predicative possession

Predicative possession is a clause where the ownership a certain object is predicated of a possessor (‘Kolbeinn has a fox’).

Stassen (2011b) made a typological survey of the expression of predicative possession, the most frequent strategy in his sample was the habeo-verb-construction (like in Germanic and Romance languages). 26% (63/240) of the languages in that study used habeo-verb strategy. It is however quite likely that this survey is biased and the amount of habeo-verb-languages overrepresented. The habeo-verb construction takes the possessor as subject and the possessum as object (Kolbeinn a un renard - ‘Kolbeinn has a fox’).

The other major types in Stassen’s study are syntactically intransitive clauses which have the basic form of an existential clause. These constructions all feature a locational or existential predicate, often translated as ‘to be at’, ‘to be there’ or simply ‘to exist’. Collectively they make up the rest of the sample, 74% of the languages in the sample.

Stassen divides these intransitive existential clause into four subtypes: 20% of the languages in Stassen’s sample has an oblique-genitive-construction, 9% oblique-locational, 20% topic and 25% a conjunctional construction. The most frequent subtype in Stassen sample is the conjunctional, where the possessor is the subject and the possessum accompanied by locational or topical marker. In the languages of his sample this marker originates from a marker of simultaneity between clauses.

The oblique locational and the topical-construction is equally common in Stassen’s sample. In the oblique locational construction the possessor NP is marked for locational relation and the possessum is the subject. The possessum is the subject in the topical construction as well, but the possessor is in this case marked as the topic. The oblique genitive possessive construction is the least common, it is similar to the locational in that the possessum is the subject but there is no locational interpretation.

In Stassen’s sample we find that Bambara is coded as having a locational construction and Hausa conjunctional. Diola-Fogny and Mòore employs habeo-verbs.

Table 18 displays the different expression of predicative possession found in the FT-material of the corpus.

<table>
<thead>
<tr>
<th></th>
<th>Total pred poss</th>
<th>gagner only</th>
<th>ya only</th>
<th>yena only</th>
<th>null only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41 (100%)</td>
<td>9 (22%)</td>
<td>18 (44%)</td>
<td>5 (12%)</td>
<td>5 (12%)</td>
</tr>
</tbody>
</table>

Table 18: Predicate possession in FT

Out of the few instances of predicative possession ya is most frequent and gagner second.

Ya and yena does function as the sole marker of predicative possession. Etymologically they both contain instances present tense 3rd person of the French habeo-verb avoir; a. That does not, however, mean that ya and yena in FT are instances of the habeo-verb.

Ya and yena can function as locative or existential statives (see section 5.4.2 och 5.4.3). There are instances of predicative possession with no obvious verbal elements whatsoever,
labeled “null” in the table. Locative and existential clauses have also been found with no verbal elements.

In Stassen’s subtypes of predictive possession with intransitive existential clauses the possessum is the grammatical subject, we do not know if this is true of the occurrences of *ya, yena* or the clauses with no verbal element. The constructions are not oblique locational, there are no locational markers on the possessor NP. As for the conjugalational subtype, there are a few instances where *ya* and *yena* could be interpreted as relativizers. The issue of topic-marking in FT is unknown.

It is very likely that the expression of predictive possession with *ya* (and to a certain extent *yena* and null) is one of Stassen’s intransitive existential subtypes, but which one is not clear. We need to know more about topic-marking and *ya* and *yena* as relativizers.

The anonymous manual suggests that the most common strategy for expression predicative possession involves the verb *gagner,* it would seem that *ya* is more common — perhaps as a result of unknown substrate influence or linguistic universal.

### 5.2.8 Copula relations

The term copula can be defined in many different ways. For the purpose of this study, a copular clause is a clause that expresses a stative relationship between two phrases, most often Nominal Phrases (NP) and Adjectival Phrases (AP), but also Prepositional Phrases (PP). This definition of copular clauses focuses primarily on the “stativeness” of the relation between the phrases. There are definitions of copular clauses which include the notion of ‘become’ - this is not done here.

The different copular clauses found in the material have been divided into five different categories of copula relations.

<table>
<thead>
<tr>
<th>Copula relation</th>
<th>example</th>
<th>formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>presentative</td>
<td><em>That’s Harry!</em></td>
<td>DEM = AP/NP</td>
</tr>
<tr>
<td>predicative</td>
<td><em>The cat is grey</em></td>
<td>NP = AP</td>
</tr>
<tr>
<td>equative25</td>
<td><em>Harry is a cat</em></td>
<td>NP = NP</td>
</tr>
<tr>
<td>locative</td>
<td><em>Harry is under the sofa</em></td>
<td>NP = PP</td>
</tr>
<tr>
<td>existence</td>
<td><em>Harry exists</em></td>
<td>NP =</td>
</tr>
</tbody>
</table>

Table 19: Copula relations distinguished in this study

It is widely known that all languages need not have an overt copula for expressing the equative (Stassen 2011c), this is often called 'zero copula'. The majority of the languages in Stassen’s study (54%, 211 out of 386) did not have the possibility for zero copula in equative clauses. Among these languages were Malgasy, Hausa, French, Mòoré, Wolof, Bambara, Maninkakan (Western) and Mandinka (Gambian).

Table 20 covers the different expressions of copula relations in the corpus.

---

25Equative copula is similar to ‘predicate nominals’ as used by (Stassen 2011c) and copula verbs indicating a Identity relation as used by Dixon (2010b:157).
Table 20: Expression of copula relations in FT

<table>
<thead>
<tr>
<th></th>
<th>Total COP</th>
<th>Total</th>
<th>zero</th>
<th>only ya</th>
<th>only yena</th>
<th>être</th>
<th>cest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>498 (100%)</td>
<td>169 (34%)</td>
<td>176 (35%)</td>
<td>87 (17%)</td>
<td>22 (4%)</td>
<td>34 (7%)</td>
<td></td>
</tr>
<tr>
<td>COP Predicative</td>
<td>272 (100%)</td>
<td>93 (34%)</td>
<td>115 (42%)</td>
<td>40 (15%)</td>
<td>10 (4%)</td>
<td>8 (3%)</td>
<td></td>
</tr>
<tr>
<td>COP Presentative</td>
<td>74 (100%)</td>
<td>26 (35%)</td>
<td>19 (26%)</td>
<td>7 (9%)</td>
<td>1 (1%)</td>
<td>20 (27%)</td>
<td></td>
</tr>
<tr>
<td>COP Equative</td>
<td>100 (100%)</td>
<td>39 (39%)</td>
<td>26 (26%)</td>
<td>20 (20%)</td>
<td>6 (6%)</td>
<td>6 (6%)</td>
<td></td>
</tr>
<tr>
<td>COP Locative</td>
<td>14 (100%)</td>
<td>3 (21%)</td>
<td>4 (29%)</td>
<td>3 (21%)</td>
<td>4 (29%)</td>
<td>6 (43%)</td>
<td></td>
</tr>
<tr>
<td>COP Existence</td>
<td>32 (100%)</td>
<td>7 (22%)</td>
<td>7 (22%)</td>
<td>17 (53%)</td>
<td>0 (0%)</td>
<td>6 (19%)</td>
<td></td>
</tr>
</tbody>
</table>

The column of être in table 20 includes être (6 occurrences), est (12), étais/t (2) (avoir/ya) été (2) and seraient (1). There are very few occurrences of être and it is likely that many of the instances are due to authors being influenced by standard French.

Zero-marking of copula relations is very frequent in FT, as the anonymous manual suggests. It is not found in the lexifier, nor in several of the potential substrate languages.

I have treated cest as a word in and of itself. Cest is etymologically combination of the demonstrative pronoun ce and the copula verb être. This decision to count cest as one non-compositional word is based on the fact that cest can occur as a sole copula verb with other subjects and that there are no occurrences of lone ce. Out of the 22 occurrences of cest, 10 are analyzed as subject-less and 11 have another explicit subject. Cest is found in the work of 25 different authors, most of the occurrences of cest are before 1900.

(22) Français Tirailleur (FrTir0966 Béchet 1889:78)

Original Ça c’est noirs comme nous autres

Gloss 3pl COP black like/as 1pl

‘They are black like us’

There are several different means of expressing copula relations in FT, we need to understand the nature of adjectives as a part-of-speech in FT and polysemous ya and yena better in order to fully understand copula in FT. The analysis by Wilson (1999) and the anonymous manual that FT lacks a copula is potentially contradicted by this material.

5.2.9 ‘Being able’ — moyen and pouvoir

There are three different constructions that are given translations of ‘be able’: moyen, pouvoir and yena manière. The last one is very minor — there is only one instance. Etymologically moyen is a noun meaning ‘means’ or ‘way’, pouvoir is a French verb equivalent to English ‘can’ and faire manière is a FT construction which in French means literally ‘do/make manner/way’ (23).
Français Tirailleur (FrTir1080, Marceau 1911:32)

Original: Toubab, yena manière tout

Stdzd: Toubab, yena manière tout
gloss: white.person yena manner all

‘The whites can do everything’

*Moyen* is much more frequent, 70 out of 79 instances of ‘be able’ are expressed. The remaining 8 cases are expressed by the verb *pouvoir* (7 *pouvoir, 1 peut*). *Pouvoir* occurs more in the older material whereas *moyen* mainly appears after 1900.

5.2.10 Summary: function seeking form

The main strategy of expression for polar interrogation appears to be through intonation, standard negation is expressed by pre-verbal *pas* and non-standard negation is very similar (or even identical) to the lexifier.

There is little reason to assume that FT has a productive grammatical gender system and nor is there a difference in natural gender in the pronoun system. There is no distinction between subject and object position of the pronouns. In order to express attributive possession FT employs possessive pronouns if the possessor is pronominal and singular. If the possessor is nominal attributive possession can be formed either with juxtaposition (possessum - possessor) or prepositional construction with *de* (but there is evidence of other preposition as well).

The most frequent verbal element in clauses which predicts possession is *ya* and secondly *gagner*. In the study by Stassen (2011b), it is common for predicative possession to be expressed by an intransitive clause involving a verb which can denote existence, this could be true of FT. The important difference lies in the marking of the possessum, which in Stassen’s analysis is the subject of the clause. We do not know if this is true of FT as well.

There is indeed zero-marking of copula-clauses, like Wilson and the anonymous manual suggests, but copula relations can also be marked by *ya* and *yena*. In order to fully understand copula relations we need to define whether or not adjectives constitutes a distinct part-of-speech.

The notion of ‘be able’ can be expressed by *moyen*.

5.3 Form seeking function

This section is devoted to the exploration of the function of certain very frequent forms in the material, the mirror image of the previous section.

5.3.1 Drop of overt subject

There are clauses that contain an explicit formal subject, but this subject carries no semantic meaning. These are called “dummy subjects” or “expletives”. The most commonly
example of this are statements about the weather, such as *it rains* and existence *there are no ghosts*. Some languages, like English require an explicit subject pronoun here even when it contributes no meaning, other languages do not require these semantically empty formal subjects.

In French it is obligatory to mark person agreement with the subject on the verb. French is coded as having 'obligatory pronouns in subject position' (Dryer 2011a). It was possible, however, to omit pronominal subjects in Old French26 (Adams 1987:2) and in Latin. There is still the possibility to omit subjects in sentence-initial position in French in certain registers and early production (Haegeman 1997:236). Studies on English (Nariyama 2004) suggest that while English is not pro-drop, it is indeed possible to omit pronominal subjects in casual conversation.

There are many instances of subject drop in the FT-material, and the phenomenon is not unique to any certain time period or author of document. Table 21 displays the overall occurrences of drop of overt subject. ‘Drop of dummy’ stands for instances where it is clear that the omitted subject falls within the category of dummy subjects. ‘Null Subject’ stands for instances where there is a semantic subject. If it is unclear in the FT clause whether there is null subject (3sg) or drop of dummy that clause has been annotated for null subject. There is a division between dynamic and stative clauses, and a further division of the stative clauses into different copula-subtypes.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>All drop</th>
<th>Null subject</th>
<th>Drop of dummy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>1 737</td>
<td>(100%)</td>
<td>318 (18%)</td>
<td>263 (14%)</td>
</tr>
<tr>
<td><strong>Indep</strong></td>
<td>1 521</td>
<td>(100%)</td>
<td>289 (19%)</td>
<td>216 (14%)</td>
</tr>
<tr>
<td><strong>Sub</strong></td>
<td>208</td>
<td>(100%)</td>
<td>25 (12%)</td>
<td>18 (9%)</td>
</tr>
<tr>
<td><strong>Dynamic</strong></td>
<td>822</td>
<td>(100%)</td>
<td>115 (14%)</td>
<td>104 (12%)</td>
</tr>
<tr>
<td><strong>All COP</strong></td>
<td>498</td>
<td>(100%)</td>
<td>127 (26%)</td>
<td>95 (19%)</td>
</tr>
<tr>
<td><strong>COP existence</strong></td>
<td>29</td>
<td>(100%)</td>
<td>22 (76%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td><strong>COP locative</strong></td>
<td>17</td>
<td>(100%)</td>
<td>2 (12%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td><strong>Only <em>ya</em></strong></td>
<td>192</td>
<td>(100%)</td>
<td>54 (28%)</td>
<td>43 (22%)</td>
</tr>
<tr>
<td><strong>Only <em>yena</em></strong></td>
<td>97</td>
<td>(100%)</td>
<td>24 (25%)</td>
<td>7 (7%)</td>
</tr>
</tbody>
</table>

Table 21: Drop of overt subject

Drop of subject is quite common in the FT-material, primarily in copula clauses. Imperative clauses often lack overt subject and have not been included in table 21, 63% of the imperative clauses had no explicit subject.

In 83 of the clauses with subject drop the subject was identical to that of a immediately preceding clause, making the task of the listener to identify the referent much easier. It is possible that there are more instances where there is contextual information available to the listener, we cannot know for sure. Here is one example of this type of subject drop.

26Old French is often defined as spoken 1100-1300.
(24) Français Tirailleur (FrTir0565, Diallo 1926:86)

**Original** Moi pas bien connaître dire français, [...] mais connais lire français dans cœur à toi

**Stdzd** Moi pas bien connaître dire français [...] mais connaît lire français dans cœur à toi

**Gloss** 1sg.NOM NEG good know speak French [...] but know read French in heart at 2sg.ACC/DAT

‘I don’t speak French well, [...] but I can read the French of your heart’

There are 10 instances of the copula *cest* where there is no overt subject, these have been coded as instances of subject drop (see section 5.2.8). If we were to analyze *cest* as containing a pronoun there would be 328 cases of drop of over subject (incl ’dummies’), i.e.19% of all clauses instead of 18%.

Which subjects are most often dropped then? If we consider the translation and surrounding clauses, it seems like most deleted subjects are of 3rd person singular. Table 22 displays the amount of explicitly marked subjects and the null subjects, i.e. drop of overt subject that is not a dummy.

<table>
<thead>
<tr>
<th></th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit subject</td>
<td>309</td>
<td>204</td>
<td>336</td>
<td>71</td>
<td>9</td>
<td>40</td>
</tr>
<tr>
<td>Null subject</td>
<td>54</td>
<td>26</td>
<td>140</td>
<td>12</td>
<td>0</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 22: Null subject distributed over person and number

We cannot be certain of the conditions in which the subject can be dropped. It could be an influence of older varieties of French or a result of the spoken nature of FT. In most contact varieties with a French lexifier in the APiCS-sample it is obligatory to have an overt pronominal subject if the referent is highly activated. (Haspelmath and the APiCS Consortium 2013)\(^{27}\). The issue of drop of overt subject is not discussed in previous literature on FT, it needs to be investigated further.

\(^{27}\)Once more, these varieties are most often defined as creoles.
5.3.2 Polysemous form *ya*

One of the most striking elements of FT is *ya*. In 1915 the company Banania used the slogan *y’a bon* to market their chocolate powder in France (figure 10). Banania’s advertisement for chocolate powder still features a West African soldier, a *Tirailleur Sénégalais*. However, today the character speaks standard French.

*Ya* occurs in 39% of the strings in the corpus, from the material from 1860’s all the way until the 1950’s. In 198 clauses (11%) there is no other obvious verbal element than *ya* (see table 23).

The origin of *ya* is the *il y a*-construction in French, which means ‘there is’. This construction has two functions: firstly existence/location (25) and secondly ‘passed time’ (26).

(25) French, *il y a*, existence (own example)

```plaintext
il y a un renard dans ton jardin
3sg.NOM there have.PRS.3sg a fox in 2sg.POSS garden
```

‘There’s a fox in your garden’

(26) French, *il y a*, passed time (own example)

```plaintext
Kolbeinn est né il y a trente-et-un ans
Kolbeinn be.PRS.3sg born 3sg.NOM there have.PRS.3sg thirty-one years
```

‘Kolbeinn was born thirty-one years ago’

Figure 8: Advertisement for chocolate powder. ©1915 Giacomo de Andreis.
The French construction must be formed with the 3rd person pronoun *il* in both instances. As we saw in section 5.3.1, it is not common (or even possible) to drop overt subjects in French. The pronoun *il* is not very common in FT overall, the most common form for the 3rd person pronoun is *lui*, out of the 204 occurrences of 3sg subject position there were only 17 of *il* (see section 5.3.5).

Most of the occurrences of *ya* in the corpus material are not formed with *il*, many even lack overt subject all together. The most common position of *ya* is directly before the predicate (27), be it verbal or adjectival. This is true of 98% of the occurrences of *ya*.

(27) Français Tirailleur (FrTir0313, Cousturier 1920:55)

**Original** quand moi y a parler ça...

**Stdzd** Quand moi ya parler ça

**Gloss** when 1sg ya speak DEM

‘when I said that...’

If there are negators (*pas, jamais, rien, plus*) they occur after *ya* and before the main verb. This is identical to the position of negators in French with respect to the finite auxiliary verbs (see section 5.3.2.).

All instances of *ya*, both when occurring before other verbs and when it appears alone, were tested for correlations with different potential relevant tense-aspect-contexts. The TA-contexts have been inferred from the context and translations. There was no clear correlation, it would seem that there are no overt markers of TA in FT.

Table 23 shows how *ya* co-occur with different functions and contexts.
The occurrence of *ya* in the corpus suggests that it is highly polysemous, but it does not seem to have a temporal or aspectual function. *Ya* does not occur alone as a dynamic verb, but precedes dynamic verbs — perhaps as a marker of finiteness as Wilson suggests. When alone *ya* has a stative meaning of copula or possession.

There is only one instance that could be interpreted as the 'passed time'-construction of French:

(28) Français Tirailleur (FrTir1180, Anon 1899:243

*Original* /.../ *moi pas content il y a vingt-deux jours*

*Stdzd* moi pas content ya vingt-deux jour

*Gloss* 1sg.NOM NEG be.happy *ya* twenty-two days

‘I wasn’t happy twenty-two days ago/It’s been twenty-two days since I was happy’

*Ya* also forms a near-obligatory/obligatory part of two verbal constructions. *ya besoin* ‘need’ and *ya moyen* ‘be able/work’. Table 24 shows the amount of strings that contain *ya* and *yena* in stative and dynamic clauses; and clauses with *besoin* ‘need’ and *moyen* ‘be able to’ present.

Table 23: *Ya* in FT

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th><em>ya</em></th>
<th>only <em>ya</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1 737 (100%)</td>
<td>682 (39%)</td>
<td>198 (11%)</td>
</tr>
<tr>
<td>Dyn</td>
<td>823 (100%)</td>
<td>297 (36%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>All stative (incl Pred. Poss &amp; COP)</td>
<td>908 (100%)</td>
<td>383 (42%)</td>
<td>197 (22%)</td>
</tr>
<tr>
<td>All COP</td>
<td>499 (100%)</td>
<td>183 (37%)</td>
<td>176 (35%)</td>
</tr>
<tr>
<td>COP Equative</td>
<td>100 (100%)</td>
<td>26 (26%)</td>
<td>26 (26%)</td>
</tr>
<tr>
<td>COP Predicative</td>
<td>272 (100%)</td>
<td>121 (44%)</td>
<td>116 (43%)</td>
</tr>
<tr>
<td>COP Presentative</td>
<td>74 (100%)</td>
<td>21 (28%)</td>
<td>19 (26%)</td>
</tr>
<tr>
<td>COP Locative</td>
<td>14 (100%)</td>
<td>4 (29%)</td>
<td>4 (29%)</td>
</tr>
<tr>
<td>COP Existence</td>
<td>32 (100%)</td>
<td>7 (22%)</td>
<td>7 (22%)</td>
</tr>
<tr>
<td>Pred. Poss.</td>
<td>40 (100%)</td>
<td>22 (55%)</td>
<td>18 (45%)</td>
</tr>
<tr>
<td>Subj: 3sg</td>
<td>476 (100%)</td>
<td>209 (44%)</td>
<td>89 (19%)</td>
</tr>
<tr>
<td>Subj: N</td>
<td>318 (100%)</td>
<td>171 (54%)</td>
<td>33 (10%)</td>
</tr>
<tr>
<td>Null subject</td>
<td>263 (100%)</td>
<td>130 (49%)</td>
<td>45 (17%)</td>
</tr>
<tr>
<td>Drop of dummy</td>
<td>55 (100%)</td>
<td>24 (44%)</td>
<td>12 (22%)</td>
</tr>
<tr>
<td>PST</td>
<td>279 (100%)</td>
<td>99 (35%)</td>
<td>8 (3%)</td>
</tr>
<tr>
<td>PROG</td>
<td>79 (100%)</td>
<td>26 (33%)</td>
<td>4 (5%)</td>
</tr>
<tr>
<td>FUT</td>
<td>113 (100%)</td>
<td>33 (29%)</td>
<td>5 (4%)</td>
</tr>
</tbody>
</table>
After examining the instances of *ya* in the anonymous manual, Wilson (1999) draws the conclusion that it is a marker of finiteness and that it is also a stative verb denoting existence. It would seem that *ya* does indeed occur mostly before the predicate, be this verbal or non-verbal, however, as can be seen in the table above, *ya* does occur alone without any other verbal elements.

In the case of predicative and presentative copulas, it is possible to argue that *ya* is not the only verbal element. If we make no distinction between adjectives and verbs, but rather view adjectives and verbs as one part-of-speech in FT, then the instances of *ya* before these adjective-verbs is not very different from *ya* before other verbs. This could be described as a predicate marker or of finiteness.

This is, however, not an appropriate analysis of the possessive and existential/locational instances though, we have no evidence that there is any other element in those clause capable of carrying the possessive or existential/locational meaning.

Previous literature has suggested that *ya* is not a copula (and that there is no overt copula in FT) but a marker of finiteness and sometimes also stative verb of location/existence (Wilson 1999). In our corpus material we find instances of *ya* as a marker of predicative possession (perhaps as a verb of location/existence, see section 5.3.7) and equative copula.

It is possible that the function of denoting location/existence is closely related to that of predicative possession. Many languages employ a intransitive existential construction, including presumably much influential Bambara.

Our corpus suggests that *ya* is highly polysemous. The following functions are attested in several different sources:

- equative copula (26 occurrences)
- stative verb predicating location or existence (11 occurrences)
- stative verb predicating possession (18 occurrences)
- before adjectives, potentially predicate marker or copula verb (116 occurrences)
- before dynamic verb, potentially as predicate marker (297 occurrences)
- obligatory part of verbal construction (109 occurrences)

### Table 24: *Ya besoin* and *ya moyen*.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th><em>ya</em></th>
<th>only <em>ya</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1 737 (100%)</td>
<td>682 (39%)</td>
<td>198 (11%)</td>
</tr>
<tr>
<td>All stative (incl Pred. Poss &amp; COP)</td>
<td>908 (100%)</td>
<td>383 (42%)</td>
<td>197 (22%)</td>
</tr>
<tr>
<td>Dynamic</td>
<td>823 (100%)</td>
<td>297 (36%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td><em>besoin</em></td>
<td>55 (100%)</td>
<td>48 (87%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td><em>moyen</em></td>
<td>78 (100%)</td>
<td>61 (78%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>
5.3.3 Polysemous form \textit{yena}

Another highly frequent element of FT is \textit{yena}. \textit{Yena} stems from the French construction \textit{il y en a} which also means ‘there is’. \textit{Il y en a} is very similar to \textit{il y a}, the important difference lies in \textit{en}. \textit{En} is a pronoun which denotes 3rd person accusative or the prepositional construction of \textit{de} + NP. The habeo-verb \textit{avoir} (\textit{il y en a}) requires an object, this means that \textit{il y en a} can be an independent clause in standard French in contrast to \textit{il y a} which is not an independent clause.

\textit{Il y en a} is a relatively uncommon construction; in a corpus of modern spoken French (Debrock et al. 2001) it occurred 564 times in a collection of 902 756 words (cf. \textit{y a}: 6201 occurrences).

Table 25 shows how \textit{yena} co-occur with different functions and contexts.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>\textit{yena}</th>
<th>only \textit{yena}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1 737 (100%)</td>
<td>249 (14%)</td>
<td>97 (6%)</td>
</tr>
<tr>
<td>Dyn</td>
<td>823 (100%)</td>
<td>103 (13%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>All stative (incl Pred. Poss and all COP)</td>
<td>908 (100%)</td>
<td>145 (16%)</td>
<td>96 (11%)</td>
</tr>
<tr>
<td>All COP</td>
<td>499 (100%)</td>
<td>93 (19%)</td>
<td>87 (17%)</td>
</tr>
<tr>
<td>COP Equative</td>
<td>100 (100%)</td>
<td>22 (22%)</td>
<td>20 (20%)</td>
</tr>
<tr>
<td>COP Predicative</td>
<td>272 (100%)</td>
<td>43 (16%)</td>
<td>40 (15%)</td>
</tr>
<tr>
<td>COP Presentative</td>
<td>74 (100%)</td>
<td>8 (11%)</td>
<td>7 (9%)</td>
</tr>
<tr>
<td>COP Locative</td>
<td>14 (100%)</td>
<td>3 (21%)</td>
<td>3 (21%)</td>
</tr>
<tr>
<td>COP Existence</td>
<td>32 (100%)</td>
<td>17 (53%)</td>
<td>17 (53%)</td>
</tr>
<tr>
<td>Pred. Poss.</td>
<td>40 (100%)</td>
<td>8 (20%)</td>
<td>5 (13%)</td>
</tr>
<tr>
<td>Subj: 3sg</td>
<td>476 (100%)</td>
<td>53 (11%)</td>
<td>30 (6%)</td>
</tr>
<tr>
<td>Subj: N</td>
<td>318 (100%)</td>
<td>105 (33%)</td>
<td>52 (16%)</td>
</tr>
<tr>
<td>PST</td>
<td>279 (100%)</td>
<td>48 (17%)</td>
<td>7 (3%)</td>
</tr>
<tr>
<td>PROG</td>
<td>79 (100%)</td>
<td>16 (20%)</td>
<td>3 (4%)</td>
</tr>
<tr>
<td>FUT</td>
<td>113 (100%)</td>
<td>19 (17%)</td>
<td>1 (1%)</td>
</tr>
</tbody>
</table>

Table 25: \textit{Yena} in FT

As table 25 shows \textit{yena} occurs often in combination with other verbal elements, but it can mark different copula relations and also predictive possession on its own. The corpus material does not seem to suggest any temporal or aspectual function of \textit{yena}, but this is needs to be further investigated.

French makes use of a habeo-verb to express predicative possession, \textit{avoir}. This verb forms part of the \textit{il y en a} that is the origin of \textit{yena}. We cannot exclude the possibly that this is the reason that \textit{yena} can mark predicative possession. It is however also possible that \textit{yena} in its function as a marker of location/existence can mark predicative possession as well. Many languages in the study by Stassen (2011b) made use of an intransitive construction involving an existential verb for expressing predicative possession (see section 5.3.7).
The anonymous manual states that *yena* is used in a periphrastic construction to express adjectival meaning (29) and suggests that this is a result of substrate influence. These examples are however only found in the anonymous manual and not in any of the other sources.

(29) Français Tirailleur (FrTir1267, Anon 1916)

Original
Tirailleur y en a pas bon, lui pas obéir

Stdzd Tirailleur yena pas bon, lui pas obéir
gloss tirailleur yena NEG good 3sg.NOM NEG obey

‘A bad soldier doesn’t obey’

This appears to be the case for *yena* as a relativizer as well, that function is almost entirely restrict to the utterances from the anonymous manual.

Our corpus suggests that *yena*, like *ya* is highly polysemous:

- as equative copula (20 occurrences)
- stative verb predicating location or existence (20 occurrences)
- before adjectives, potentially predicate marker or copula verb (40 occurrences)

*Yena* is similar to *ya* but less frequent. The use of *yena* in predicate possession might, like *ya* be an instances of Stassen’s intransitive existential possession constructions, but there are too few examples.

The issue of *yena* as a relativizer needs to be further investigated.

5.3.4 *Gagner* — to get, become or have?

In Standard French *gagner* means ‘win’, ‘gain’, ‘acquire’ or ‘get’. There is in total 65 occurrences of *gagner* in the corpus. In the majority of these occurrences *gagner* appears to be very similar to the etymological lexical form in French. 11 cases have been translated as ‘become’ and 9 as ‘have’.

Most of the cases of *gagner* as ‘have’ are found in the anonymous manual, suggesting that it is nor necessarily true for all FT. The ‘become’-meaning however is found in several sources, all of them after 1916 (31).

(30) Français Tirailleur (FrTir0138, Marie-Victoria 1921:499)

Original Toi, ma sœur, ya gagné Commandant

Stdzd Toi, ma sœur, ya gagner Commandant
gloss 2sg.NOM my.sister ya become boss

‘You sister, you shall be the boss’
The function of *gagner* appears to be subject to change and not restricted to possession.

### 5.3.5 Wanting, liking and being happy — *content* and *vouloir*

There are two strategies for expressing ‘want’ in FT; *content* (18 occurrences) and *vouloir* (29). There are quite few occurrences, but it seems as if *vouloir* is older occurring primarily before 1900 and *content* in the sense of ‘want’ is more recent.

Etymologically, *content* is an adjective meaning ‘satisfied/happy’ and *content* occurs 20 times in the sense of ‘happy and 20 as the verb ‘like’. One might expect the sense of ‘enjoy/like’ to be more recent, but (31) dates to 1899.

(31) Français Tirailleur (FrTir0660, Montrozier 1902:155)

**Original**

Mahomet pas content

**Stdzd** Mahomet pas content

**gloss** Mohamet NEG like/enjoy/be.happy

‘Mohamet wouldn’t like that’

Like *gagner*, *content* seems to be subject to change and perhaps pushing out the older form of *vouloir* in the function of ‘want’. Further investigation into function of *content* as a verb meaning ‘be happy’ might shed light on the statue of adjectives as a part-of-speech in FT.

### 5.3.6 Reduplication

Reduplication is defined by Rubino (2011) as the “repetition of phonological material within a word for semantic or grammatical purposes”. Languages can employ full reduplication (the repetition of words or roots) but it is also possible to have partial reduplication (the repetition of smaller segments such as phonemes).

In order for a language to be classified as having reduplication this process needs to be productive, i.e. applicable to a set of open class words and still in current use. In Rubino’s sample 277 languages have partial reduplication, 35 full reduplication and 56 no productive reduplication.

There are many possible functions of reduplication, Rubino lists a plural number, continued aspect and case among many others.

Corne (1999:201) suggests that there is reduplication in FT and that it marks intensity or continuity. Rubino lists intensity and continuity as possible functions of reduplication. There is no mention of reduplication in the thesis on FT by Wilson.

It has been claimed that reduplication is a common phenomenon in pidgins, as it is in many creoles. Bakker (2003) challenged that perception and successfully proved that reduplication is not as common in pidgins as it had previously been described.
While there is not productive reduplication in French, there is productive reduplication in some of the substrate languages - most notably there is full and partial reduplication in Hausa, Maninkakan (Western) and Mòoré. There is also reduplication in Wolof, but only of full words or roots.

Let us examine what evidence of reduplication there is in the corpus. If we consider all instances where there is any form of repetition of a preceding element (be it an entire word or a single phoneme) we end up with 35 potential candidates of reduplication. 12 of these are found in the anonymous manual from 1916.

The first examples that we can exclude is the onomatopoetic *tam-tam* found in Fr-Tir0395a (Binger 1892:120-1) and FrTir0395a (Claudel-Hubin 1991).

Secondly, there are instances which are loans from the West African languages and that we should exclude from evidence of productive reduplication (table 26).

- *gris-gris* amulet (Anon 1915:152)
- *grigri* medicine (Béchet 1889:86)
- *bagabaga* termites (Béchet 1889)
- *ménéméné* black ants (Béchet 1889:78)

Table 26: Loans with reduplication

Thirdly, we have the issue that reduplication is defined as occurring within a word. However if we trust the segmentation of material into words as done by the authors the remaining instances of reduplication are not word-internal. The concept of a *word* is central in distinguishing 'reduplication' from 'word iteration/repetition'.

Gil (2005:33) discusses this problem and notes that the repetition of words can have similar functions to reduplication, such as intensity of iterative aspect. Gil has set up a set of criterions for distinguishing reduplication from word repetition/iteration. Unfortunately several of these criteria cannot be applied to the FT-material since they rely on phonological distinctions and we do not have access to spoken data. Many of the instances of potential reduplication fulfill criteria of both Gil’s repetition and reduplication.

According to Gil it is very unusual, but not unheard of, that reduplication produces more than 2 copies. In the FT material we have three instances of three copies. These are most likely instances of word repetition. Another criteria that Gil proposes is that the input in the reduplication process should be one word (or less), which means that we can rule out *Femme tout nu, tout nu!* (Nordenck 1886:276) since it is one of the clear examples of multiple word-input.

12 of the remaining 28 instances are found in the anonymous manual. All of them express with spaces in between the words and what can be best described as intensification (32) or continuity.
Now, it is not possible to say with great certainty that these instances are reduplication or word repetition/iteration that marks intensity/emphasis/continuity. However, these forms are very infrequent, mainly restricted to the anonymous manual and since it is possible for word repetition to have similar function as reduplication the meaning is not crucial to the distinction.

It would appear that FT does not have productive reduplication, it is at least not as very frequent feature.

5.3.7 Summary: form seeking function

Subjects are sometimes omitted in FT, 18% of the material. In many of these instances (83 to be precise) the subject was the same as in the preceding clause, suggesting that it is possible to omit subject under identity (what is known as SS in switch-reference languages). We cannot know the nature of subject drop in FT for certain, we have too little information. The literature suggests that it could be influence from French or a feature of casual spoken interaction.

Ya has many functions, one of these may be as a predicate marker preceding dynamic verbs and a stative verb denoting possession, location/existence and equative copula. In order to fully understand ya we need to understand adjectives as a part-of-speech — if it is not distinct from verbs then ya could be analyzed as a predicate marker in more instances.

Yena has previously been analyzed as a stative verb and relativized. Yena as relativized is not attested in enough sources. The function of yena is, like ya, dependent on our definition of adjectives.

Gagner ‘win/acquire/get’ is often used in FT in the sense of ‘get’ or ‘have’, but in more recent material also as ‘become’.

Content in FT can mean ‘want’, ‘like’ or ‘be happy’. The lexifier form of vouloir for expressing ‘want’ is more common in older material, it would seem that both gagner and content are subject to change.

One can argue that reduplication is present in FT, but it is not a frequent feature and is much more frequent in the anonymous manual than any other source.
5.4 Notes on diachrony

5.4.1 Order of modifier and head

Another potential trend is the order of modifier and head within a Noun Phrase. There are slightly more occurrences of the head preceding the modifier in more recent material. However, the construction with a postposed numeral that we find in the anonymous manual seems to be restricted to the manual. In Bambara and many other West African languages, the numeral is postposed to the noun (Dryer 2011f). It might be that the author of the anonymous manual has exaggerated the influence of Bambara in FT.

5.4.2 Aspectual fini?

There is one instance of fini in an unusual position (33) that might indicate a potential strategy for marking completive, perfective or perfect aspect.

(33) Français Tirailleur (FrTir0506b, Maran 1921:96)
Original Moi y’en a croire lui crêvé fini,
Stdzd Moi yena croire lui crêver fini,
gloss 1sg.NOM yena believe 3sg.ACC die finish

‘I think he has passed out’

There are other examples of fini in the sense of ‘finish doing/being’, but this example is different. Judging from the post-verbal position of fini, the context and translation it would seem that it is possible to interpret it as aspectual.

Bybee et al. (1994:67) have noted that one possible origin for markers of Resultativity, Perfective or Perfect aspect is indeed verbs of ‘finish’. In a typological survey of the Perfect by Dahl and Velupillai (2011) 21 languages (20% of all that had Perfect) had a marker that was derived from a word meaning ‘finish or ‘already’. There is also evidence of fin(i) as marker of completive aspect in several French creoles (Holm 1988:162-163). There is also a post-verbal construction of the verb ‘finish’ together with the infinite marker that can mark completive aspect in Bambara (Bird et al. 1977 as cited in Holm 1988:163). If there is a completive/perfective/perfect construction in FT which originates from a verb meaning ‘finish’ it could be a result of a linguistic universal or substrate influence.

Example 33 is however the only example of this potentially aspectual fini in FT.

5.4.3 Pre-verbal qui/que

There is one form that shows up in the early material (34), pre-verbal qui/que (labeled KV in the standardized material). Other occurrences of qui/que appear, as in the lexiifer, clause-initially\(^{28}\), but this is not the case here.

\(^{28}\)This means that the adjectival periphrastic constructions as suggested by the anonymous manual are most likely excluded.
(34) Français Tirailleur (FrTir0553, Nordenck 1886:286)

Original Moi qu’a conni un capitaine

Stdzd Moi kV connu un capitaine

Gloss 1sg.NOM kV know un capitaine

‘I know/knew a captain’

There 15 occurrences of this, 13 of these are from the 1880’s, suggesting that this was perhaps an early construction that didn’t last. It has been noted by seven different authors. The production of the most recent one is dated to 1931 (35).

(35) Français Tirailleur (FrTir0352, Laverrière 1932:147)

Original Père, vite, la chapelle qui brûle

Stdzd Père, vite, la chapelle kV brûle

Gloss father fast DEF chapel kV burning

‘Father, quick, the chapel is on fire!’

This form always occur before the predicate and it is not found more often in clauses that have a past tense-reading. The form might be similar in function to ya and yena. There is a possibility that the form marks focus, but this is purely speculative.

5.4.4 Summary: notes on diachrony

It can be hard to distinguish between actual historical trends in FT and authors who are more less influenced by French.

There are few potential authentic trends visible in the corpus, the use of moyen for ‘be able’ as opposed to pouvoir (section 5.2.9), content as ‘want’ instead of vouloir (section 5.3.5) and the order of modifier and noun (section 5.4.1).

The issue of aspectual fini (section 5.4.2) can only be speculation at this point, but does have interesting theoretical implications. As for the pre-verbal qui/que it is possible that it is a marker of focus or predicate, but it would seem that the form disappeared early on.

The order or elements, not only modifier and head, but perhaps also of the elements of the finite clause would be an interesting issue to pursue further.
6 Discussion

There are many more features of FT that should be further investigated (see section 3.2.3), and we are, as always, in need of more data nor only on FT but also on pidgins in general.

One reoccurring issue is the spoken nature of FT, it was rarely written. We need to know more about the spoken French of that time in order to fully understand FT. The conditions that play into spoken conversation must be taken into account when working with pidgins. Omission of subject is one such example where the spoken modality could be relevant, when speaking it is possible to make use of many other means of transmitting meaning (shared knowledge, pointing, etc). These additional means of communication is not possible in the written form to the same extent.

We expect pidgins to have a restricted lexicon and many polysemous words, which could be reflected in a low TTR. We also expect spoken material to have a lower TTR than written. The standardized TTR of this corpus is 27%, is this necessarily unusually low? More comparisons between pidgins and spoken data of languages lacking inflectional morphology is needed.

There are many things we do not yet know about spoken conversation. As our knowledge of spoken language grows, so will our understanding of the mechanisms of pidginization.

It has been stated that FT lacks overt copula (Anon 1916 and Corne 1999). It is not clear what exactly they mean by 'copula', I assume that they primarily refer to equative, presentative and predicative copula relations (excluding locative and existential). In order to answer this question we must first understand ya and yena and if there is indeed a distinct part-of-speech of adjectives.

In the case of ya it is possible to analyze it as a marker of finiteness or the predicate instead of copula verbs29. This, in combination with the high number of zero forms of copula in predicative copulas, brings up the issue of whether or not it is relevant to make a distinction between adjectives and verbs in FT.

An argument for the lumping together of these two categories would be the lack of verbal morphology and syntax, there are few formal differences between (what is etymologically) adjectives when they occur in predicative position and what is defined as verbs in French. They can both be preceded by ya or yena.

We could define adjectives in FT as 'verby' when they occur in predicate position, using the terminology of Wetzer (1996). In the typological study of the expression of predicative adjectives Stassen (2011a) found that verbal encoding was more common than non-verbal. In Wolof what can be defined as predicative adjectives are expressed verbally, Mandinka, Môrë and Bambara has both strategies Stassen (2011a). Verbal encoding of what is often adjectives in IE is also common in creoles (Holm 1988:176-177).

An argument for the division of verbs and adjectives into two different parts-of-speech would be that words that are etymologically adjectives can occur within an NP, verbs cannot. A second argument for the division is the occurrence of être and cest, if adjectives

29This is not an exhaustive definition of ya and yena, but it is one potential explanation of their function.
and verbs belong to the same category why can they be preceded by another verbal element?

Another argument against lumping verbs and adjectives into the same part-of-speech is that adjectives can modify the head of an NPs in FT, and in that they do differ from verbs. The status of adjectives as a part-of-speech in pidgins need to be further investigated if we are to understand the grammar of FT in general and the nature of ya and yena in particular.

Many of the predictions by the anonymous manual and Delafosse are supported by the corpus material. However, there are features that appear exclusively in the manual, such as post-nominal numerals and the periphrastic adjectival constructions with yena. These features are both similar to structures that exist in Bambara and Wolof, two languages that the author(s) of the manual most likely knew were spoken by the soldiers. The author(s) of the manual might have exaggerated Bambara’s influence (Van Den Avenne (2012:258) even calls the FT of the manual a calque of Bambara). There is also the possibility that these features occurred more often in the variety of FT that the manual portrays.

The results of this study suggests that the influence from West African substrate languages on these features is minor, it would seem that the overall reduction of the pidginization is the more important. It has been suggested that structures that are typologically common are more likely to show up in reduced language varieties such as pidgins. Examples of these structures are: the expression of polar interrogation through intonation and pre-verbal negation. The reason why they show up in pidgins might be due to the fact that (i) they are common in the world’s languages (i.e. more likely to occur in the involved languages), or (ii) they are “cognitively less complex” and therefore preferable in these contact situations. We need more empirical investigations of pidgin corpuses to further understand this.

It is worth noting that much of the new material of FT that hasn’t been brought up previously in the literature does behave like the manual. The structures of FT that have been described in this thesis are found in the majority of the sources, this suggests that FT was conventionalized enough to be labeled a ‘pidgin’.

63
7 Conclusions

In this corpus study we can discern the following of the nature of FT:

- polar interrogation is, most likely, expressed through intonation (see section 5.2.1)
- standard negation is expressed through pre-verbal particle pas (see section 5.2.2)
- non-standard negation is expressed identical or very similar to the lexifier French (see section 5.2.3)
- there is no productive system of grammatical gender (see section 5.2.4)
- there is, most likely, no distinction of natural gender in 3rd person pronouns (see section 5.2.4 and 5.2.5)
- the pronoun system consists primarily of: moi, toi, lui, nous, vous and eux (see section 5.2.5)
- when the possessor is pronominal attributive possession is expressed by possessive pronouns (see section 5.2.6)
- when the possessor is nominal attributive possession is expressed through juxtaposition and prepositional constructions (see section 5.2.6)
- when attributive possession is expressed through juxtaposition the relative order is possessum - possessor
- there is a a highly polysemous morpheme ya that can occur as marker of predicative possession, existential/locational verb, equative copula and possible as a predicate marker and/or predicative copula (see section 5.3.2)
- there is a second a highly polysemous morpheme yena that can occur as a existen- tial/locational verb and potentially also as a copula (see section 5.3.3)

There are instances of potential change over time, but nothing can be said for certain as for the diachronic development of FT (see section 5.4).

It is very likely that comparisons with research on spoken language can contribute tremendously to our understanding of the lexicon and structure of pidgins and pidginization, in particular corpuses of spoken utterances of languages lacking inflectional morphology. We need more empirical studies of pidgins to understand their nature.

The features of FT that were found in this thesis are documented in the works of many different authors. This suggests that FT did indeed exist and that it was stable enough for us to label it a ‘pidgin’. As is the case with all studies of language, there is great internal variation and the different language varieties are merely overlapping in enough many areas for linguists to be satisfied with calling it one entity.
References


Anon (1899). Par ici, par là. *Journal de la maison, 9.*


65


Leymaire, H. (1898). *La mort de Mabiala le féticheur. À travers le monde*.


