

Established French loanwords in Sango – A pilot study

Christina Thornell

1. Introduction

This study is an initial attempt to describe the established French loanwords in the lingua franca Sango, which is the main language spoken in the Central African Republic (CAR). It is a second language for the majority of the population of 3 million people for whom a great number of vernaculars serve as first languages. French loanwords in Sango are prompted by the contact between Sango and French which has occurred over the past century. Loanwords form an integral part of the recipient language. In the present study, the loanwords integrated both with regard to form and meaning will be focused on. These are characterized by stability, both diachronic and synchronic, and linguistic adaptation. Diachronic stability is suggested by, for instance, Haugen 1950 as being an important criterion for the establishment of loanwords, and synchronic stability by Weinreich 1968, whereas linguistic adaptation is a widely recognized criterion for loanwords. The established loanwords will further be described with regard to word classes, semantic domains and functions.

The present study is based on part of a tape-recorded corpus. The original corpus consists of approximately 65 hours of tape-recorded speech, representing various discourse types. 120 informants have contributed to the corpus. The recordings were made in 1990-1991 in the western part of the CAR. For the present purpose, I restrict the study to approximately 4,000 running items involving 500 running words from interviews and conversations with four women between 19-33, with different socio-linguistic backgrounds. For the diachronic aspect of loanwords, sources from 1908 to 1978 have been consulted. Furthermore, observations are made on the basis of my total stay of eleven years, working with among other things literacy and book publishing in the CAR, and my knowledge of Sango acquired during this period.

Before commencing the description of loanwords, some sociolinguistic and linguistic aspects of Sango will be presented to explain the motivations for French loanwords in Sango

2. Background

2.1 Sociolinguistic aspects

The CAR is a multilingual society. French was introduced by the Whites and serves by tradition as the official language. However, since March 1991, Sango is promoted as an official language alongside French. French represents the most prestigious language and proficiency in it is a prerequisite for social advancement. A relatively restricted sector of the population, however, shows a high proficiency in the language, which they have acquired within the formal education system. The great majority of the population shows a limited proficiency or none at all. One of the many vernaculars functions, in general, as an L1, but for an increasing number of young people in urban areas, Sango serves as the L1.

2.2 Origin and status of the Sango language

Sango is a recently arisen language, having developed out of the Ngbandi languages. The Ngbandi languages form part of the Ubangian subgroup of the Niger-Congo languages, as do the majority of the languages in the CAR. The exact time of Sango's coming into being is currently being discussed. Diki-Kidiri 1982 and Pénel 1984 suggest that it existed before the arrival of the Europeans in the latter part of the nineteenth century. Samarin 1989, on the other hand, argues that it came into being during a short period of time some years after the establishment of the Europeans in the Oubangui area, and that it arose between the indigenous people and Africans from other places who were employed by the Europeans. Regardless of the time of its coming into being, it is, however, commonly agreed that the activities of the Europeans led to its becoming more widespread.

The status of Sango is also discussed in various works. Samarin 1967 and 1982 argue that it is a creolized language, as does Pasch 1993 based on linguistic criteria. Diki-Kidiri 1987, on the other hand describes Sango as being at one of the end-points of the Ngbandi continuum. This does not necessarily assign Sango to pidgin or creole status. The geographical proximity of the lexifier languages indicates that Sango may not be characterized as a typical pidgin/creole because of the required criterion of the geographical distance from its lexifier language.

2.3 The standardization of Sango

The promotion of Sango first as a national language and then as a language of official status requires its standardization orthographically, grammatically and lexically. Institutions of language cultivation are concerned with these issues. The most important one is *l'Institut de Linguistique Appliquée (ILA)*, founded in 1975, pursuing the aims of *la Commission Nationale Centrafricaine pour l'Etude de la Langue Sango*.

Orthographical rules. Sango is primarily an oral language and has been used in writing almost exclusively by religious organizations. This has led to various spelling systems, the most important of which were developed by an American Protestant denomination and the Catholic Church. The decree *Code de l'orthographe officielle du Sango* in 1984 by President André Kolingba forms a serious attempt to unify the different existing orthographies. For more details of the different spellings in Sango and the standardized orthography, see Thornell 1994.

As French loanwords are in focus in the present study, the phonemic system of Sango and French and their phonotactic pattern will briefly be compared. The Sango phonemes stated in the decree are to a great extent overlapping with the French ones, but not completely, e.g. the French phoneme / \emptyset / is lacking in Sango. This phoneme in French loanwords in Sango is replaced by [e], as in [pn \emptyset] 'tyre', which becomes [pine:], spelled *pinee*. In addition, there are differences in Sango and French phonotactic patterns according to Diki-Kidiri 1977, Bouquiaux, Kobozo & Diki-Kidiri 1978 and *LETAC* (1983-84). They argue that the phonotactic pattern in Sango is composed of two syllable types, namely V, and CV. This pattern should also include French incorporations regardless of their original phonotactic structures, as is illustrated in [pine:]. The pronunciation of French incorporations in accordance with the standardized norms in Sango when this conflicts with the French norms is, however, not widespread. The pronunciation seems to be a social marker. This is evident from my study of French incorporations in Sango in Boyd, Andersson & Thornell 1991, which is based on the tape-recorded speech of twelve male multilingual speakers. This study shows that phonological integration does not seem to be the norm for speakers proficient in French. Pronunciation with regard to the speaker's proficiency in French reveals that the pronunciation of French incorporations tends to be a function of education, in that the speakers of little formal education had the highest rate of phonological adaptation to Sango. These speakers are in addition the oldest of the speakers, which may

in part also contribute to the adaptation. The application of the standardized rules seems to be restricted to the few individuals who are conscious of the standardization process and agree with it, but these speakers limit the application of these norms to formal speech.

Sango is a tone language. The 1984 Sango Orthography Decree distinguishes three distinct tones; low, middle and high, and three modulated tones; low-high, high-low and mid-low. These are marked as follows:

<i>low tone:</i>	unmarked
<i>middle tone:</i>	˘
<i>high tone:</i>	ˆ
<i>modulated tones:</i>	as composed tones

At present, the tone marking is restricted to items that only differ with regard to tones.

Lexical standardization. The focus of the standardization, besides being on the orthography of Sango, has been and is still on the vocabulary, which among other things has resulted in dictionaries. The first dictionary at the official level is Bouquiaux et al. 1978. Some of the wordlists compiled earlier have played a certain role in the compilation of this dictionary, such as Tisserant 1950. The dictionary is, however, largely based on tape-recorded material representing various parts of the CAR, predominantly the central and eastern part. The dictionary, besides indigenous items, also includes loanwords and neologisms referring to concepts related to modern society. Further dictionaries of official character are provided by *LETAC* (1983-84), which is a result of the five-year-long crosslinguistic lexicon project *Lexiques thématiques de l'Afrique Centrale*, involving most of the African French-speaking countries. This project was aimed at lexical modernization within the economic, social and educational areas. The two means available in this project were either to have recourse to borrowing or to use the resources of Sango. Primarily, the resources of Sango are taken into consideration which resulted in neologisms but quite a lot of borrowings are also accepted. Most of the borrowings originate from indigenous languages, such as the Ngbandi languages. However, French loanwords seem to be avoided.

An emerging standardized variety. Bouquiaux et al. 1978 and *LETAC*, in contrast to other dictionaries compiled earlier, can be considered to constitute the norm for the vocabulary, as they were compiled by direction

from official institutions. This norm is in the present study termed the standardized variety and characterized, as already mentioned, by less tolerance of French items, by a high degree of newly coined items, and by borrowings from indigenous vernaculars. *LETAC* is extreme at this point whereas Bouquiaux et al. 1978 allows more French items. The standardized variety is used in news broadcasts, and in official speech by a limited number of individuals, such as those people involved in institutions of language cultivation and students at the university level for whom courses in Sango form part of their training. In addition, some individuals have been influenced by this variety by listening to radio programmes. The standardized variety is, apart from some oral contexts, applied in the few official writings that are published, e.g. *Diki-Kidiri* 1983 and the reading-book for the first grade *Use buku ti Kossi* published in 1981.

The development of a standardized norm as presented above is reflected in Bible translations according to my study of some Sango versions of the Gospel according to St. Mark. The 1966 translation, which was published before norms were explicitly formulated by institutions of language cultivation, comprises more French items than the 1986 translation, in which the directions from ILA, the principal institution of language cultivation, are applied. In the 1966 translation the lexical frequency is 15.1%, or 115 items. This number of French items seems rather normal with respect to the non-indigenous cultural context to which the text refers. The occurrence of French items in the 1986 translation is extremely low – 1.0% of the total number of the words, or 24 items.

A general variety. In contrast to the standardized variety as presented above is the variety that in the present study will be termed the general variety. This variety is characterized, among other things, by a higher tolerance of French loanwords. It encompasses all writings and speech before 1970 as the serious engineering of the vocabulary of Sango began in the 1970s. It also encompasses the writings and speech that are not obviously influenced by the standardized norm. For instance daily speech reflects this variety. Thus, the present corpus represents the general variety.

2.4 Typology of Sango

Sango belongs to the isolating language type. Its word order shows a consistent SVO order which indicates the function of different lexical constituents. Thus, the word order in Sango overlaps to a great extent with

French word order. The affixes in Sango are restricted to a marker of agreement, a plural marker, and a derivational affix, e.g.:

A-mbeni kete molenge a-sala ngia na gbe tí keke
 PL-some small child AGR-make play at underside of tree
 'Some little children play under the tree.'

The marker of agreement also involves French verbs regardless of their degree of establishment.

Plural marking is restricted to the NP but it is not obligatory on either the noun or the adjective. There is, however, a tendency to mark animate nouns versus inanimate ones, according to Samarin 1967. Samarin finds that marking is avoided when the number of the referents is known. These observations are also supported in my study of plural marking reported in Boyd et al. 1991. A further observation in my study is that marking on French nominals is more frequent than on the same category of Sango nominals in the informant group as a whole, as well as with regard to the subgroups. The group that most consistently marks both French and Sango nominals on the whole are the youngest speakers, all male students. This marking also includes those nominals that are not usually marked according to Samarin. This indicates that a change might be in progress, making plural marking more consistent than before. It is probable that this change is influenced by French.

The derivational suffix *-ngö* can be applied to all verbs except for the copula. The derived items function either as nouns or as adjectives.

2.5 The Sango vocabulary

Sango is a newly arisen language, as mentioned earlier. This impacts on the size of its vocabulary. Not much lexicographical research has been done in Sango meaning that discussions of the size of the vocabulary must be based on rough estimations. Bouquiaux et al. 1978 contains 6,000 entries and constitutes the most extensive dictionary. However, this dictionary reveals considerable gaps. So the vocabulary in Sango encompasses more than 6,000 items but certainly considerably less items than the vocabulary of a standardized European language, such as French. This assumption is based on the different characters of these languages.

The great bulk of the Sango vocabulary is claimed to originate from Ngbandi. Samarin 1961, focusing on the core vocabulary of 100 items as defined by Swadesh finds that 78% of the items are Ngbandi items. Other African languages, Central African as well as non-Central African, and

European languages have contributed to the Sango lexicon. This is, for instance, observed by Tisserant 1950, Kêrux & San-Youen 1950 and Bouquiaux et al. 1978. The latter finds that among the loanwords originating from indigenous languages, by far the most frequent ones come from the Banda languages, and among loanwords originating from non-indigenous African languages, those originating from Lingala prevail. Among the loanwords originating from the European languages, the French ones predominate, as expected. Taber 1964, which is a study of French loanwords in Sango based on Samarin's fieldwork in 1962, finds that as much as 51% of the total 998 lexemes are French items. However, the textual frequency is restricted to 6.8%. Most of these French incorporations occur once or twice, meaning that these are characterized as nonce borrowings. Only a little more than ten occur frequently. These are assigned both to open and closed word classes.

2.6 Word classes

The nouns and verbs form the most basic word classes in Sango. The adjectives embrace few items and most members also function as adverbs. The form classes *prepositions* and *connectors* are very restricted with regard to single units. Most of the prepositions and connectors in Sango have a composite structure based on nouns. Their degree of grammaticalization is not clear meaning that these complex items in the present study are analysed into their components for word class assignment. The multi-functionality of the word classes is striking.

3. Assumptions on established French loanwords

Section 2 above shows that several factors favour the adoption of French items in Sango. Important factors involve the intensity of contact between the two languages, similarities in typologies, the poor morphology of Sango and the differences in the two vocabularies. The adopted items are classed as established loanwords according to the diachronic and synchronic stability criteria proposed in the introduction. The loanwords are classed on the basis of either one or both of the criteria. Of the two, diachronic stability appears to be the most important criterion, as this can be defined as synchronic stability at several points in time. Synchronic stability comprises consistent semantic functions and consistent usage patterns including the frequency criterion. The third criterion, linguistic adaptation, plays a minor role compared with the stability criteria. The pronunciation and the crucial

morphological marking of Sango affixes cannot serve as criteria for establishment, as is obvious from sections 2.3 and 2.4, whereas syntactic and semantic adaptation signal establishment.

Most of the incorporations can be described as established loanwords in the general variety but they are recognized to a lesser extent in the standardized variety. The established loanwords are to a great extent assigned to nouns and verbs but they also represent other word classes such as adjectives and connectors. Semantically, the content words, in particular the nouns, denote concepts typical of Western civilization. With regard to semantic functions, the loanwords comprise the sole item for a certain concept, but also synonyms, both 'near' absolute synonyms and cognitive synonyms. Absolute synonyms are identical except in form. Cognitive synonyms make up items that differ with regard to for instance variety.

4. Data and method

The entire corpus comprises 65 hours of Sango speech produced by approximately 120 informants from various sociolinguistic backgrounds. This speech also comprises French incorporations and a few Gbaya incorporations. For more details see next section. The corpus was recorded in 1990-1991 in the Western part of the CAR. The corpus involves both monologues, such as sermons, storytelling, and broadcasts, and dialogues, such as interviews and conversations. The speech represents both planned and unplanned speech. The sermons and broadcasts can be described as planned speech. The unplanned speech is, however, not easily associated with a specific speech type. Even many conversations seem to be planned, at least to some extent, which was probably a result of the recording-situation.

For the present pilot study, 4,000 running items produced by 4 female informants were excerpted from the corpus – 500 running items representing conversation and interview for each informant. This speech represents the general variety. The informants, aged 19-33, were students at a household management school and, in addition, they belonged to the same social network. They belong to different subgroups of the Gbaya group. They learnt Sango as well as their ethnic language as children. Three of them regarded Sango as their L1 – either alone or together with another language. The women's proficiency in French differed. Two women had a minimum of French whereas the remaining two had greater proficiency.

The Sango speech is, as far as possible, phonemically transcribed according to the spelling rules prescribed in the 1984 Sango Orthography

Decree. Bouquiaux et al. 1978 are also consulted. The French incorporations are, however, transcribed in French orthography and the few occurring Gbaya incorporations according to Gbaya spelling rules. The lexical tones are not marked because the computer programs used for processing did not allow such signs. For the transcriptions the conventions proposed by MacWhinney 1991 are used. The corpus is tagged with regard to language and word class. The French items are tagged with regard to their word class assignment in Sango. For the tagging, the PC BETA elaborated by Brodda 1991 is applied.

The analysis focuses on the French single words and NPs that are established in Sango. Their establishment is characterized by diachronic and synchronic stability and linguistic adaptation. The diachronic stability is studied based on sources from six points of time between 1908 and 1984. These sources include both the general and standardized variety. The sources Giraud 1908, Gerard 1930, Tisserant 1950 and Taber 1964 represent the general variety whereas Bouquiaux et al. 1978 and *LETAC* represent the standardized variety. The established French loanwords may either be considered to be established in both varieties or in the general variety. Synchronic stability is characteristic of the French incorporations to various degrees meaning that the French incorporations with regard to this criterion can be described along a continuum, with the end-points synchronically stable and synchronically instable. The third criterion, linguistic adaptation, comprises mainly morphosyntactic and semantic aspects. A synthesis of the three criteria is presented in next section.

5. Results

An overview of the total number of incorporations will be given. The incorporations will then be restricted to the French single items and NPs. These French incorporations will be characterized in terms of their degree of establishment in Sango. The established loanwords will then be described further with regard to word classes, semantic domains and semantic functions.

5.1 Overview of the corpus

In the present Sango corpus, comprising slightly more than 4,000 running items, approximately 87% of the textual frequency is made up of Sango items. The remaining 13% comprises French and Gbaya incorporations and proper names. The French incorporations make up about 9%, the Gbaya

Examples	The general variety	The standardized variety
<i>mais</i> 'but'	Established loanwords 107 items (78.7%)	Established loanwords 36 items (26.5%)
<i>encore</i> 'another'		
<i>bon</i> 'well'		
<i>mademoiselle</i> 'miss, unmarried woman'		↑ Unestablished loanwords 100 items (73.5%)
<i>voyager</i> 'travel'	Inbetween 22 items (16.1%)	↓
<i>rappeler</i> 'remind'	Unestablished 7 items (5.2%)	

The degree of greyness indicates the degree of establishment.

Figure 1. Model of the establishment of French incorporations in Sango.

items less than 2%, and proper names a little more than 2%. The Gbaya items and proper names will not be taken into consideration.

The great majority of the French incorporations involve single items and NPs. This includes 136 single items and NPs with a total frequency of occurrence of 315 times.

5.2 The establishment of French incorporations in Sango

The synthesis of the three criteria of establishment, diachronic and synchronic stability, and linguistic adaptation is presented in Figure 1. The figure shows that the 136 French incorporations are classed as established loanwords to a varying extent in the two varieties. The general variety recognizes the great majority, 78.7% or 107 items, whereas the standardized variety recognizes a minor part, namely 26.5% or 36 items. Those considered to be established in the standardized variety are considered to be established in the general variety as well. The great majority of the established loanwords common to the two varieties are characterized both by diachronic and synchronic stability and in some cases

by linguistic adaptation. These items will not easily undergo change with regard to their establishment. The established loanwords that are restricted to the general variety, on the other hand, are to a lesser extent characterized by the two stability criteria. Consequently, a great deal of the established loanwords in this variety may be subject to change with regard to their establishment – in either direction.

The French incorporations that are not recognized as established in the standardized variety are considered to be unestablished. These make up the majority. This contrasts with the general variety, in which an insignificant number are unestablished loanwords. These items are not all newly incorporated items but also items that are on their way out of Sango, such as *rappeler* 'remind'. A further contrast with the standardized variety is that the general variety also comprises French incorporations that can neither be classed as established nor as unestablished. These French incorporations may undergo change either in the direction of establishment or non-establishment. An example of the former is, probably, *moment* 'time' that, probably, will become established in that it is on its way to being lexicalized in the expression *na moment tî* [at time of] 'when' and in *na moment nî* [at time DEF.ART] 'at a certain point of time'. The former functions as a connector and the latter as an adverb. An example of an item that, probably, is on its way out of the language is *pêcheur* 'fisher'. This item will possibly be replaced by *wäfängö susu* [human.being.killing fish] or *zo tî fängö susu* [human.being of killing fish].

5.3 The established French loanwords with regard to word classes

The established French loanwords are shown in Table 1 according to word classes. The greatest part of the loanwords is made up of open word classes, the closed word classes are represented by few items or none at all. Among the word classes, nouns prevail. However, the established loanwords in the classes connectors, adverbs, and adjectives, although few in number, comprise all the French incorporations assigned to these word classes in the present corpus – with some exceptions for each word class.

This pattern of established loanwords reflects to some extent the character of the open and closed word classes, in general, in Sango. Furthermore, it reflects gaps in Sango, arising in contact with French. The loanwords assigned to adverbs and connectors are to a great extent accepted by the two varieties. This pattern signals that the adverbs and connectors really form an integral part of the Sango vocabulary.

Table 1. The established French loanwords with regard to word classes.

Word class	Number
Open word classes:	98
Nouns	74
Verbs	14 (15)*
Adjectives	5
Adverbs	5 (6)
Closed word classes:	6 (9)
Pronouns	0 (1)
Numerals	-
Connector	5 (7)
Prepositions	0 (1)
Interjections	1
Remainder	3**
<i>Total</i>	107

*() includes the items assigned to the remainder category.

**the verbal expression *il faut* 'one must', the multifunctional *jusqu'à* 'as far as', and *depuis* 'since'.

5.4 Loanwords designating non-indigenous cultural concepts

Established loanwords assigned to the open word classes tend in general to designate non-indigenous cultural concepts. In the present study, this holds for more than two thirds of the established French loanwords assigned to the nouns and verbs as is shown in Table 2. Looking at the two word classes separately, the tendency varies considerably. Almost 80% of the nouns represent non-indigenous cultural concepts but only 20% of the verbs do so. The non-indigenous nouns represent concepts typical of Western civilisation. The semantic domains that these nouns represent include education, time, artifacts, religion, person, the political domain, and rural life. The ranked order of these semantic domains, however, reflects to a great extent the focus of the topics of the interactions. The few nouns referring to indigenous or general concepts form part of several semantic domains but not the semantic domains mentioned above. These nouns comprise, e.g. *papillon* 'butterfly', *milieu* 'middle', and *place* 'place'.

The established French verbs designating non-indigenous cultural concepts are restricted to three items. The major part of the verbs involve grammatical verbs, e.g. *commencer* 'begin' and mental verbs, e.g. *penser* 'think'.

Table 2. Established French loanwords designating non-indigenous concepts.

	Non-indigenous		Total number
	Examples	Number	
<i>Nouns</i>		59 (79.9%)	74
Semantic domains			
Education	<i>école</i> 'school'	21	22
Time	<i>lundi</i> 'Monday'	11	12
Artifacts	<i>cartouche</i> 'cartridge'	8	9
Religion	<i>église</i> 'church'	7	8
Person	<i>portugais</i> 'Portuguese'	4	7
The political domain	<i>district</i> 'district'	3	3
Rural life	<i>café</i> 'coffee'	2	2
The remainder domain	<i>service</i> 'duty'	3	11
<i>Verbs</i>	<i>repasser</i> 'iron'	3 (20.0%)	15
<i>Total</i>		62 (69.6%)	89

5.5 The functions of the established French loanwords

The functions of an item involve many aspects. Here, the functions with regard to the organization of the vocabulary will be considered, in terms of the function of established French loanwords: as sole names for certain concepts, as synonyms, absolute as well as cognitive, and as hyponyms. The established loanwords studied are those assigned to nouns and verbs. See Table 3. Most of the items can be described by more than one function. Here, the most striking function will be taken into account.

Among the functions, synonyms prevail. This function characterizes approximately half of the established French nouns and verbs. Sole items for certain concepts represent a slightly lower number, whereas hyponyms comprise a considerably smaller number than the other functions. The number is unexpectedly low. This is probably because of the restriction of the function to French nouns and verbs that have a clear superordinate that often is used as a synonym.

Among the synonyms, the cognitive ones dominate. *Juillet* 'July' as well as the other terms designating months belongs to this category. Its cognitive synonym is *lengua* 'July'. Most of the Sango equivalents are restricted to the standardized variety. This contrasts with the French items forming part of absolute synonyms. These word pairs occur in the general variety, such as *peut* 'can' and *lingbi* 'can'. The French items forming part of absolute synonyms make up a rather great number. This is surprising as absolute

Table 3. The functions of the established French nouns and verbs.

Functions	Number	Proportion
Sole name of concept	35	39.3%
Synonyms:	44	49.5%
'Absolute' synonyms	19	21.4%
Cognitive synonyms	25	28.1%
Hyponyms	10	11.2%
Total	89	100.0%

synonyms rarely occur in the world's languages. It is questionable whether these loanwords are truly absolute, but it is obvious that they are closer to absolute synonyms than to cognitive ones. The existence of these absolute synonyms can be explained by the fact that these French items were introduced long before their corresponding Sango items and became established in their place. It is, however, possible as time goes by that some of these French nouns and verbs will lose their stability.

The established French loanwords functioning as sole items for certain concepts comprise two categories, namely the French loanwords that actually make up the sole terms for certain concepts, and those having Sango equivalents that are not known in the Western part of the CAR. An example of the former is *gâteau* 'cake' and of the latter *manquer* 'miss', which has the synonym *luku* 'miss' (unknown in the Western CAR). The Sango equivalents are dictated by the standardized variety.

The established loanwords functioning as hyponyms differentiate a semantic field. In general, their superordinates are composed of Sango items. *Bible* 'Bible' is a clear instance of a hyponym. Its superordinate is *mbeti ti Nzapa* 'book of God', which refers to the Bible as well as to all theological literature.

The functions as defined above are largely related to the definition of synonyms. This definition requires that an item is semantically and syntactically interchangeable within a certain context. This general definition is weak when studying synonyms forming part of two so different languages in their mode of expression as Sango and French. This can be illustrated by the established French verb *fatiguer* 'be tired' in Sango. In accordance with the definition above, this verb has no synonym. However, there exists an equivalent Sango expression, namely *terê ti zo a woko* [body of human.being PERS.PR become.tender] 'one is tired'. Some established loanwords indicated in the present study as having no synonyms have

semantically equivalent expressions but with completely different structures. Besides contributing to the expansion of the vocabulary, loanwords of this type involve structural change in Sango.

6. Conclusion

The present study describes established French loanwords in Sango based on a restricted tape-recorded corpus (4,000 running items) representing the general variety, which is the variety of Sango commonly used in the CAR. The loanwords are characterized by a stability criterion, involving both a diachronic and synchronic dimension. The former dimension appears to be the most crucial one. An additional criterion is that of linguistic adaptation. This criterion seems to play a minor role, at least as concerns adaptation to Sango pronunciation and morphology. The French incorporations are to a great extent composed of single items and NPs. These can in the general variety be described along a continuum with the end-points being established loanwords and non-established loanwords. Most of the incorporations can be characterized as established whereas a minor part is recognized in the emerging standardized variety according to dictionaries published by institutions of language cultivation. The loanwords recognized in the standardized variety are also accepted in the general variety. The low tolerance for French loanwords in the standardized variety, can, in part, be explained by the influence of nationalism and by influences from French puristic language ideologies. The latter is motivated by the formal education in the CAR being based on the French education system.

With regard to word classes, the great majority of the established loanwords are assigned to open word classes, in particular nouns, whereas an insignificant number are assigned to closed word classes, in particular connectors. The distribution of word classes is expected, as nouns and verbs are basic in the world's languages, and as nouns in Sango, in addition, form part of closed word classes, such as prepositions and connectors. The indigenous connectors are very restricted in number in Sango, which reveals gaps in contact with French which is rich in connectors. The difference motivates incorporation of French connectors in Sango. The incorporated French connectors are to a great extent accepted in the two varieties. With regard to word classes, the order of the particular loanword with regard to the general organization of the word classes remains to be studied. Loanwords are probably integrated with respect to a certain order.

The established loanwords are further described with regard to semantic domains and functions. The description of semantic domains is restricted to nouns and verbs. The majority of the established nouns reflect Western concepts, whereas verbs reflect such concepts to a lesser extent. This difference between nouns and verbs has to be studied further. Complete semantic domains seem to be borrowed, such as the semantic domain based on the Western time concept, the concept of formal education, and the concept of Christianity. This is quite normal as these domains were introduced by Western people via French.

With regard to functions, the most crucial function, in general, is that of gap-fillers. Gap-fillers imply that the term represents a concept for which there is no earlier term. The present study reveals that this function is important but not the most crucial. The most crucial function comprises the one of synonyms, including both absolute and cognitive synonyms. These have Sango equivalents. The absolute Sango synonyms occur in the general variety whereas the cognitive Sango equivalents are restricted to the standardized variety. The French synonyms seem to be introduced earlier than their indigenous equivalents. This is striking with regard to the cognitive Sango equivalents which to a great extent are made up of neologisms and borrowings from Ngbandi varieties. Their French cognitive synonyms refer mostly to Western concepts. The cognitive synonyms appear to be a rather recent phenomenon, which indicates that different varieties in Sango are developing. Absolute synonyms in a language are very uncommon. Therefore, the rather high number of these in Sango is surprising. They can possibly be explained by the fact that Sango is not quite standardized yet. Furthermore, it is not taught in school.

Loanwords may reinforce weak structures in the host language. This occurs when loanwords represent word classes that are restricted both in number and use, as loanwords seem to retain their subcategorization rules at least with regard to some items. With regard to Sango, this implies for instance that structures that relate events, i.e., structures involving connectors, are reinforced. Furthermore, loanwords may trigger change in syntactic structures in the host language, which is obvious from the present study. This aspect of loanwords has to be studied further. If the loanwords consistently retain their subcategorization rules, this would signal that the lexicon is prior to syntax.

The present description of established French loanwords in Sango is based on very restricted material. In some places it is indicated that the

phenomenon in question needs further investigation. However, all the tendencies revealed have to be supported by more evidence based on a larger corpus. This corpus would, in addition to speech representing the general variety, also embrace speech representing the standardized variety. In addition to a description of loanwords as presented above, further aspects of loanwords have to be added to get a more complete view of loanwords, such as the loanwords included in the standardized variety based on speech and not only on dictionaries published by institutions of language cultivation. The two varieties should possibly be regarded as a continuum rather than two distinct varieties as presented in the present study. In this wider perspective, pragmatic and sociolinguistic aspects would also be involved.

Acknowledgement

The present paper has been written as a part of my doctoral studies which are financed by The Swedish Council for Research in Humanities and Social Sciences. I am indebted to Åke Viberg for valuable comments on earlier versions of this paper.

References

- Bouquiaux, Luc, Jean-Marie Koboza & Marcel Diki-Kidiri. 1978. *Dictionnaire sango-français*. Paris: Société d'études linguistiques et anthropologiques de France.
- Boyd, Sally, Paula Andersson & Christina Thornell. 1991. 'Patterns of incorporations of lexemes in language contact: language typology or sociolinguistics', In *Papers for the symposium on code-switching in bilingual societies, theory, significance and perspective 2*, 463-488. Strasbourg: ESF Scientific Networks.
- Brodda, Benny. 1991. 'Doing corpus work with PC Beta; or, how to be your own computational linguist'. In Stig Johansson & Anna-Brita Stenström (eds.), *English computer corpora*. Berlin: Mouton de Gruyter.
- Decret no 84025 1984 *Fixant l'orthographe officielle du Sango (langue nationale centrafricaine)*
- Diki-Kidiri, Marcel. 1977. *Le sango s'écrit aussi*. Paris: Société d'études linguistiques et anthropologiques.
- Diki-Kidiri, Marcel. 1982. 'L'expansion du sango en Centrafrique'. In Jean-Pierre Caprile (ed.), *LACITO-documents Afrique 8*, 28-42. Paris: Société d'études linguistiques et anthropologiques de France.

- Diki-Kidiri, Marcel. 1983. *Kua ti kodoro*. Paris: Société d'études linguistiques et anthropologiques.
- Diki-Kidiri, Marcel. 1987. 'La situation du sango en République Centrafricaine'. In Henri Boyer & Pierre Dumont (eds.), *Etudes de Linguistique Appliquée* 65, 102-109. Didier Erudition.
- Gerard, M. 1930. *Sango, langue commerciale de l'Oubangui-Chari*. Rome: la Sodalité de S. Pierre Claver.
- Giraud, Gaston. 1908. 'Vocabulaire des dialectes sango, balkongo et azande'. *Revue Coloniale, mai 1908*, 263-91. Paris: Challamel.
- Goundjika, Mathias M. (ed.). 1981. *Use buku ti Kossi*. Bangui: L'Institut Pédagogique National.
- Haugen, Einar. 1950. 'The analysis of linguistic borrowing'. *Language* 26, 210-231.
- Kêrux, E. & San-Youen, 1950. *Sango*. Chambéry: TRP, Basile de Belleaux.
- LETAC = *Lexiques thématiques de l'Afrique Centrale: 1 Activités économiques et sociales* (1983), *2 Lexiques thématiques de l'Afrique Centrale 2, Education et formation* (1984). Paris: ACCT et Yaoundé: CERDOTOLA
- MacWhinney, Brian. 1991. *The Childes project: Tools for analyzing talk*. Hillsdale, New Jersey: Erlbaum.
- Pasch, Helma. 1993. 'Phonological similarities between Sango and its base language. Is sango a pidgin/creole or a koiné'. In S. Mufwene Saliko & Lioba Moshí (eds.), *Topics in African linguistics*, 279-292. Amsterdam: John Benjamins.
- Pénel, Jean-Dominique. 1984. 'Ethnies'. In Pierre Vennetier (ed.), *Les atlas jeune afrique-Republique Centrafricaine*, 24-25. Paris: Les éditions jeune afrique.
- Samarin, William. 1961. 'The vocabulary of Sango'. *Word* 17, 16-22.
- Samarin, William. 1967. *A grammar of Sango*. The Hague: Mouton.
- Samarin, William. 1982. 'Colonization and pidginization on the Ubangi River'. *Journal of African Languages and Linguistics* 4, 1-42.
- Samarin, William. 1989. *The black man's burden, African colonial labor on the Congo and Ubangi Rivers, 1880-1900*. Boulder, Colorado: Westview Press.
- Taber, Charles R. 1964. *French loan words in Sango: A statistical analysis of incidences*. Unpublished M.A. thesis. Hartford: Hartford seminary.

- Thornell, Christina. 1994. 'Reflections on the 1984 orthography decree'. In Helma Pasch (ed.), *The national official language of the Central African Republic. Procedures of the colloquium "The status and uses of Sango in the Central African Republic"*, 33-57. Köln: Rüdiger Köppe Verlag.
- Tisserant, R. 1950. *Sango-Oubangui-Chari*. Issy-les-Moulineaux: Les Presses Missionnaires.
- Weinreich, Uriel. 1968. *Languages in contact*. The Hague: Mouton.