CONTRASTIVE ANALYSIS IN MULTILINGUAL SOCIETIES: A METHODOLOGICAL ISSUE

Moses Kwadwo Kambou
Université de Ouagadougou, Burkina Faso
kambou.moses@univ-ouaga.bf

In language-contact situations we know that various languages influence each other. In order to best understand the degree of influence and predict the difficulties of the learners, the contrastive analysis approach was proposed in the 1940s by Fries (1945) and his disciples. Applied contrastive analysis as advocated by Lado (1957) should take into consideration the differences and similarities of the languages being compared. In the developments that followed in this area of study, there has been a lot of research in monolingual societies and the monolingual model has become the norm. In this paper I will show that the monolingual model is not appropriate as an applied contrastive linguistic model for multilingual societies and would therefore propose a multilingual model that takes into consideration the intermediary languages between the native language and the foreign language (English in the case of Burkina Faso). This model is inspired by error analysis, which takes into consideration all the languages present in the language learning situation. I will illustrate my proposed model with examples of contact between English and Mooré in Burkina Faso and then between French and Twi in Ghana.

I. Introduction

This cross-linguistic study is to determine if the monolingual model of the contrastive analysis (CA) approach can be used to predict difficulty of Mooré-French or Twi-English bilinguals in learning English or French, respectively. The study will be limited to the noun-phrase structure of the languages concerned. The hypothesis that underlies this study is that the monolingual model of CA cannot be used to correctly predict L2 learner difficulty and errors in multilingual contexts. The theoretical framework for the CA is that of structural linguistics.

Mooré is a Gur language, a subgroup of the Niger-Congo family classified in Greenberg 1966 as I.A.3 Voltaic. This language is spoken mainly in Burkina Faso and Ivory Coast, West Africa. Twi is a dialect of Akan, and is a Kwa language, a subgroup of the Congo-Kordofanian family classified in Greenberg 1966 as I.A.4 Kwa group. It is spoken in Ghana, West Africa.

I chose to examine the methodological issue of CA because students' dissertations I have read or for which I was a member of the defense committee revealed to me that most of the students use the monolingual model for their comparison, although they claim the results of their study should help improve the
teaching of English in Burkina Faso. In this respect, their studies miss certain generalizations regarding the place of the intermediary language, French, in their predictions.

Another reason for choosing this topic is that most of the maîtrise students in the Modern Languages Department of the University of Ouagadougou do CA for their dissertations. These research works are pedagogically oriented, hence I believe they should understand not only the difference between applied and theoretical CA, but also the methodological differences between monolingual and multilingual contexts. This is because in most of the CA studies, their aim is to discover and predict learning difficulties of English as a Foreign Language (EFL) students by comparing the native language with the foreign language. In this endeavor, they tend to forget the intermediary language because most of the papers they read and rely on concern monolingual contexts: English-French, Spanish-English, Polish-English, Romanian-English, etc. They are therefore heavily influenced by the monolingual model, which, as I will show later, makes wrong predictions in multilingual contexts, hence is inappropriate. This paper seeks to call their attention and also that of CA researchers in multilingual societies to the fact that the methodology they are used to is not suitable for their research context.

It is an established fact that in monolingual societies we compare L₁ with L₂ or native language (NL) with the target language (TL) because the L₂ is the only interfering language. My research question then is: How should the model for multilingual societies be? The acquisition model can help explain the CA model used simply because in a monolingual situation we have the native language compared with the target language, hence the only interference possible is from the language being learned; the languages in the contact situation are the L₁ and L₂. Can this model be valid for multilingual situations? I do not think so, since only two languages in the monolingual context are taken into consideration by the model. In the multilingual situation, we have at least three languages; this is the very reason why I believe that the multilingual situation has to have a model that takes into consideration all the languages in the language-contact situation as is the case in the monolingual context. This model could consist of two stages because the number of comparison is equal to the number of languages minus one (i.e., number of languages - 1). Considering this logic for comparison, we have only one comparison in a monolingual society where we have two languages and two comparisons in a multilingual society with three languages. Hence, our multilingual CA model is as follows:

a) compare the intermediate language with the foreign/target language:
   \[ L_{\text{int}} \rightarrow L_3. \]

b) compare the native language with the foreign/target language:
   \[ L_1 \rightarrow L_3. \]

In this model, \( L_{\text{int}} \) represents one of the various European languages (English, French, Portuguese, Spanish) that the learner has learnt after the native language, \( L_1 \), before coming to learn the foreign language (\( L_3 \)). My model as I conceive it will be valid for Sub-Saharan Africa where English, French, Spanish, and Portu-
guense are second or foreign languages, depending on the colonizer’s language. In this model, $L_1$ is the native language(s) of the learner, and $L_{int}$ is French, English, or Portuguese depending on the learner’s colonizing country.

Data for this study are from two sources: Kambou 1992 for Mooré and introspection for the Twi data. In the data, Twi and Mooré are native languages ($L_1$). French is $L_{int}$ in Burkina and $L_3$ in Ghana while English is $L_{int}$ in Ghana and $L_3$ in Burkina. The analysis of the data is based on comparing the two models with the same data in order to show the weaknesses of the monolingual model in multilingual contexts. The preliminary results of this study show that the multilingual model makes better predictions than the monolingual model for multilingual contexts.

2. Review of the literature

Comparison of languages is as old as linguistics itself. A comparison could be interlingual (comparing two or more languages) or intralingual (comparing dialects of the same language). In general, when dialects or languages are compared the aim is to show the differences and similarities between them. Thus, historical linguistics used the comparative method for the reconstruction of languages in history. Language teaching in the 40’s contributed in encouraging linguists who undertook research in contrastive linguistics in order to improve, if possible, existing language teaching methods. At that time, most people thought the main problem in language learning was the structural difference between the source language ($L_1$) and the target language ($L_2$). The pioneers of contrastive analysis applied to language pedagogy are: Fries (1945) and Reed, Lado, & Shen (1948). Research by Fries & Pike (1949) as well as that of Weinrich (1953) was focused on the learning of foreign languages. In the pedagogic perspective of contrastive linguistics, there are as many models of grammar as there are theoretical models of description. Among these, we can mention: (a) the structuralist or ‘taxonomic’ model elaborated from research by Bloomfield (1933), Fries (1945), Lado (1957), and Harris (1963); (b) the case-grammar model elaborated from Fillmore’s (1968) research on universal semantics; (c) Krzeszowki’s (1974, 1976) model based on a bi-directional contrastive studies. Most of the research on the pedagogic contribution of contrastive linguistics of the 60’s compared English, on the one hand, with German, Italian, Spanish, or French, on the other hand. The majority of these focused on phonetics and phonology; there is relatively little research on syntax. Proponents of the pedagogic perspective of contrastive linguistics in the 60’s believed that thanks to contrastive studies (of $L_1$ and $L_2$) focused on the differences, they could predict learners’ difficulty and errors. The hypothesis that underlies these researchers’ works is known in the literature as the Contrastive Analysis Hypothesis (CAH). The importance given to the phenomenon of language transfer in foreign language learning at the time brought about three versions of this hypothesis: (a) a strong version supported by Lado (1957); (b) a weak version supported by Wardhaugh (1970); and (c) a moderate one supported by Oller & Ziahosseiny (1970). Today, contrastive linguistics has gone beyond structural comparison and is using a sociolinguistic approach, because structural and semantic equivalence
are not sufficient. We must look at pragmatics and conversational implicature as was suggested by Y. Kachru (1976). S. N. Sridhar (1980) also pointed out some of the above mentioned problems in CA and showed that in spite of these difficulties, CA remains both theoretically and methodologically the most principled component of a theory of errors. Even though UG in SLA does not make it explicit, the parametric variation view of White (1988) and the parameter-setting model of Flynn (1987) are a type of CA approach to SLA.

CA, as Fisiak (1981:1) noted, 'comparative studies in linguistics have a long history.' He makes a distinction between comparative historical linguistics, comparative typological linguistics, and contrastive analysis or contrastive study. He further showed that the last two share some common elements and could be considered as synchronic comparative linguistics. 'Contrastive linguistics', according to Fisiak, 'may be roughly defined as a sub-discipline of linguistics concerned with the comparison of two or more languages or subsystems in order to determine the differences and similarities between them.'

In the literature, CA is said to have two aspects or sub-areas: applied and theoretical. In theoretical CA, the comparison between languages A and B seeks to find out how a universal category X is realized in the two languages. There is no question of directionality in this approach. The question of directionality is important to applied CA whose aim is to find out how a universal concept X realized in language A as Y, is realized in language B. The results of applied CA are said in the literature to be used for teaching, translation, bilingual studies, etc.

The review of the literature shows that the CA model used in European countries compared all the languages in the contact situation, which are generally two. There is no published research to my knowledge that questions the general CA model used in multilingual societies. There is also no published CA data in multilingual countries taking into consideration all the languages in the contact situation; it is this vacuum that my paper sets to fill and hence contribute data to the literature on multilingual contexts and also propose a multilingual CA model.

3. Data for CA

The data for the CA are from English, French, Mooré, and Twi. They are presented as L₁ (Mooré and Twi), L₂ (French and English), and L₃ (English and French). The L₂ and L₃ languages depend on the context because English is L₂ in Sub-Sahara Anglophone countries and French L₃; in Sub-Sahara Francophone countries French is L₂ and English L₃.

In this section, we simply present sentences in the three languages of the Anglophone and Francophone situations. The comparison proper will be done in the discussion section. In column I, we have the Francophone data, and in column II, that of the Anglophone situation. The noun phrases (NPs) are underlined in the examples. The NPs in the sentences or alone illustrate cases where nouns are non-specified, unspecified, or specified.

The structure of the African languages is Noun + Determiner and that of the European languages is Determiner + Noun as in examples (i) and (ii).
I. MOORE, FRENCH, ENGLISH

i. L₁: págá wá
   woman +sg/det/
   ‘the woman’
L₂: La femme
L₃: The woman

A. NON-SPECIFIED NOUNS (GENERIC USE)

1. L₁: Bàag tarà nào a nàasè.
   /dog-sg/have+inacc-aff/leg+pl/four/
   ‘A dog has four legs’
L₂: Un chien a quatre pattes.
L₃: A dog has four legs.
2. L₁: B rátà péterollè.
   /they/need-inacc-aff/kerosine/
   ‘They need kerosine’
L₂: Ils ont besoin de pétrole.
L₃: They need kerosene.
3. L₁: Biisim sao bééré!
   /milk/be better-aff/beer/
   ‘Milk is better than beeter’
L₂: Le lait est meilleur que la bière.
L₃: Milk is better than beer.

II. TWI, ENGLISH, FRENCH

ii. L₁: ñbàa ñu
   woman +sg/det/
   ‘the woman’
L₂: The woman
L₃: La femme

1. L₁: ñkramáiwɔ nni ñni ɛnai
   /dog-sg/have/leg+pl/four/
   ‘A dog has four legs’
L₂: A dog has four legs.
L₃: Un chien a quatre pattes.
2. L₁: wɔ mɔ hiá kresì
   /they/need/kerosine/
   ‘They need kerosine’
L₂: They need kerosene.
L₃: Ils ont besoin de pétrole.
3. L₁: mirik ye cè bia
   /milk/be better/beer/
   ‘Milk is better than beer’
L₂: Milk is better than beer.
L₃: Le lait est meilleur que la bière.

B. UNSPECIFIED NOUN (INDEFINITE USE)

4. L₁: M pa tari lìqd ye.
   /l/neg1/have/money/neg2/
   ‘I have no money’
L₂: Je n’ai pas d’argent.
L₃: I have no money.
5. L₁: F tàrà bòbò bìì?
   /you/have-inacc-aff/sweets/interro/
   ‘Do you have any sweets?’
L₂: As-tu des bonbons?
L₃: Have you any sweets?
/book/blue/pl/  
‘Blue books’  
L₂: Des livres bleus.  
L₃: Blue books.

7. L₁: Wà-y né sè és mà bá!  
/come with+imp/chair/pl/  
‘Bring some chairs’  
L₂: Apportez quelques chaises!  
L₃: Bring some chairs!

8. L₁: Sébr sen ya bòl.  
/book-sg/relative pro/be+inacc/blue/  
‘A blue book’  
L₂: Un livre bleu.  

/Seydu/have+inacc-aff/cat+pl/many/  
‘Seydu has many cats’  
L₂: Seidou a beaucoup de chats.  
L₃: Seidou has many cats.

10. L₁: Bi-púgla tárà líg d zèmbülá.  
/child-girl+sg-det/have+inacc-aff/  
‘The girl has enough money’  
money/  
L₂: La fille a assez d’argent.  
L₃: The girl has enough money.

11. L₁: B rátà kóom wòsgó.  
/they/need+inacc-aff/water+sg/much/  
‘They need much water’  
L₂: Ils ont besoin de beaucoup d’eau.  
L₃: They need much water.

C  
SPECIFIED NOUNS (DEFINITE USE)

12. L₁: A Sayiíd kàremmdà a sbérè.  
/Seidu/read-inacc-aff/his/book-sg/  
‘Seidu is reading his book’  
L₂: Seidou lit son livre.  
L₃: Seidou is reading his book.

/Hawa/read-inacc-aff/her/book-sg/  
‘Hawa is reading her book’  
L₂: Hawa lit son livre.  
L₃: Hawa is reading her book.
/ he/eat-acc-aff/meat-det/all /
‘He has eaten all the meat’
L₂: Il a mangé toute la viande.
L₃: He has eaten all the meat.

15. L₁: Pipi karèmbúisa.  
/ first/read-child+pl-det/ 
‘the first students’
L₂: Les premiers élèves.
L₃: The first students.

4. Discussion

Before beginning the discussion we should first of all remind the reader that the European countries at the advent of CA were monolingual countries hence the name of the CA model used in these countries. On the contrary, most of Sub-Saharan countries are multilingual due to the contact of African languages and that of their colonizers (English, French, Spanish, Portuguese) and the urge today to learn foreign languages, those mentioned above as well as German, Russian, Italian, etc.

The problem with multilingual societies is that CA is not just comparing L₁ and L₂ but we have to take into consideration the intermediary languages. For example, if we consider the case of a Francophone country like Burkina Faso, French is an intermediary language and the national or Burkinabè languages are the L₁s of the learners. What is certain is that both the native language(s) and French have an influence on the learning of English by the Burkinabè students. This condition is also true of Ganaian students learning French, but here the intermediary language is English and the native language(s) the L₁. This reality should be taken into consideration if we wish to use the results of our research to help improve the teaching and learning of English and French as foreign languages in Burkina Faso and Ghana, respectively. As is mentioned in the literature, CA alone cannot be used to improve teaching of a language since the factors affecting language acquisition and learning are linguistic, sociolinguistic, and psycholinguistic, hence it is just one of the main factors. The criticisms leveled against CA also show that CA needs the support of other approaches to language learning to sufficiently address the issues of language acquisition and learning.

The data from Anglophone and Francophone countries of Sub-Saharan Africa will be tested against the two models in competition: the monolingual CA model and the multilingual CA model.

**MONOLINGUAL CA MODEL STAGE OF COMPARISON**

L₁ → L₂  
stage 1

stage 2

**MULTILINGUAL CA MODEL**

a. L₁ → L₃

b. L₁ → L₃
4.1 Predicting the presence or absence of a determiner

From the two models above, we see that in the multilingual model there are three stages in the comparison while in the monolingual model there is only one stage. We could therefore assert that the monolingual model is a subset of the multilingual model, similar to stage (c). If this model is therefore used as the model for multilingual contexts, it will miss the information contained in stages (a) and (b) of the multilingual model. Let us analyze the data with the two models:

In examples (1-3), the monolingual model, comparing L_1 and L_3, will predict that the Mooré speaker will omit the determiner of the NP because in the L_1 there is no determiner. Hence the prediction is that the speaker will produce the following sentences in English: *’Dog has four legs’; ‘They need kerosene’; ‘Milk is better than beer’. Comparing these sentences with the expected ones in L_3 of (1-3), this shows that the prediction will be correct in examples (2) and (3). In the case of the Twi speaker, the model will predict that the speaker will also omit the determiner and say *’Chien a quatre pattes’; ‘Ils ont besoin de pétrole’; *’Lait est meilleur que bière’. Here again, if we compare the predicted forms with the L_3 expected forms, the prediction will be correct only for example (2).

As for the multilingual model, a 2-stage comparison, it will not only make the same predictions as the monolingual model, it will in addition make predictions by comparing the intermediary language (L_{int}) with the target language (L_3). Hence, it will predict that the Mooré speaker will either omit the determiner due the L_1 or use a determiner due to the L_{int} in (1). In (2), it will predict that there will be no determiner whether the L_1 or L_{int} NP is transferred into L_3, and in (3) the determiner will be dropped due to L_1 or not dropped due to L_{int}. This second model as we can see from the analysis of the data has a higher chance of making correct predictions in the three examples because it takes into consideration all the languages present in the contact situation. In the case of the Twi speaker, this model will predict that in example (1) the learner will either omit the determiner of the NP in L_3 due to L_1 or use it due to L_{int}. The L_3 sentences will be either *’Chien a quatre pattes’ or ‘Un chien a quatre pattes’. In examples (2) and (3), the Twi learner of French will omit the determiner in French due to the absence of a determiner in the NPs of L_1 and L_{int}. The speaker is thus predicted to produce the following sentences: ‘Ils ont besoin de pétrole’ or ‘Lait est meilleur que bière’. In sum, the multilingual model from the examples (1-3) will make a better prediction than the monolingual model (66% for Twi speaker and 100% correct predictions for Mooré speaker, as against 33% correct predictions for both learners).

In examples (4-11), the monolingual model will predict the absence of determiners in (4, 5, 7, 8), and the presence of determiners in (6, 9, 10, 11), due to the L_1 of the learner. In the case of the multilingual model, the prediction will be the absence of determiners in (4), due to the L_1 and L_{int} while it will predict the presence of determiners in (5, 7, 8), due to the L_{int}. In all 8 examples, the monolingual model will make 4 correct predictions (50%), while the multilingual model will make 7 correct predictions (87.5%). With regard to the Twi speaker, the monolingual model will predict the absence of determiners in the L_3 NPs in exam-
ples (4, 5, 7, 8), and the presence of determiners in examples (6, 9, 10, 11). The multilingual model on the other hand will predict the presence of determiners in the NPs of (4, 5, 7, 8) due to L_2 and in (6, 9, 10, 11) due to L_1 or L_{inf}. I can conclude that the monolingual model will make 5 correct predictions out of 8, i.e., 62.5% in the case of the Twi speaker, and 4 out of 8, i.e., 50%? in the case of the Mooré speaker. On the contrary, the multilingual model will make 7 out of 8 correct predictions, i.e., 87.5% for the Mooré speaker and 8 correct predictions out of 8, i.e., 100% correct predictions in the case of the Twi speaker.

Concerning examples (12-15), the monolingual model will predict the presence of determiners in L_3 in all the examples and the multilingual model will also predict the presence of determiners in all the examples for the Mooré speaker. As for the Twi speaker, the monolingual model will predict the presence of determiners in examples (12, 13, 14), and the absence of determiners in (15) while the multilingual model will predict determiners in all four examples. In this category of nouns, the monolingual model will make 3 out of 4 correct predictions, i.e., 75% for the Twi speaker, and 4 correct predictions out of 4, i.e., 100%, for the Mooré speaker. The multilingual model will make no wrong predictions for native speakers of either Twi or Mooré.

In conclusion, it has been shown that between the two CA models in competition, it is the multilingual model that makes fewer wrong predictions than the monolingual model. In addition, it also takes into consideration the languages present in the language-contact situation when making predictions.

4.2 Predicting the type of determiner
If the CA model in our example predicts the presence of a determiner or determiners in the target language (L_3), the next thing will be to predict the type of determiner that the L_3 learner will use. This aspect of CA needs the support of other language-learning methodologies to be very accurate. As can be seen in examples (1-3), the prediction will be that in L_3 the determiner of the NP will be the indefinite article in (1) and the definite article in (3). As regards examples (4-11), the predictions of the negative determiner ‘no’ in (4) is not possible, the prediction will be that the learner will use the normal negation, which is wrong. In (5), the prediction will be ‘some’ due to French (L_{inf}) and problems with English as L_{inf}. The case of (6) will be word order ‘some books blue’ due to French and ‘bleus livres’ due to English. In (7), ‘some’ and ‘quelques’ will be predicted correctly and the same is true for (8) and (10). However, in (9) and (11), correct predictions of the determiner to be used cannot be made since the choice of the determiner type depends on the nature of the noun in L_3, which is not the case in L_1. This is where a conceptual model of learning would be of great help. The same is true for (5) and (6).

In the third category of nouns, predicting the exact determiner will not be a problem in examples (14) and (15); however, in (12) and (13) it will be difficult, since the choice of the possessive adjective depends on the thing possessed in French, on the possessor in English, and neutral in L_1. A conceptual model will here again be useful in making correct predictions. In spite of these, there is the question of
the word order of the NP structure, which the monolingual model will predict to be noun + determiner due to \( L_1 \), and the multilingual model will predict to be Determiner + Noun due to \( L_{\text{int}} \). However, where we have adjectives and not simply articles with nouns as in (6) and (8), both models will make wrong predictions about the word order.

4.3 Predicting the source of error

In all cases the error predicted will be due to the \( L_1 \) according to the monolingual model whereas the multilingual model predicts some errors to be due to \( L_1 \) and others due to \( L_{\text{int}} \). However, in some cases, the error predicted by this model could be due to either of the previous languages learned by the \( L_3 \) learner, and in this case it is not easy to predict the exact source of the error.

5. Conclusion

The study shows that the monolingual model when used in multilingual contexts misses some generalizations with regard to \( L_{\text{int}} - L_3 \) relationship. The multilingual CA model, which takes into consideration the realities of the language-contact situation, is further given independent evidence from Error Analysis in multilingual situations, which can be considered as a complementary approach to CA.

6. Implications

There are two possible implications of this study: (a) the generalization of the multilingual CA model, and (b) the reanalysis of multilingual society data based on the monolingual CA model. The multilingual CA model as presented in this study can be generalized to Sub-Saharan Africa. For example, the case of Ghana can be generalized to all Sub-Saharan Anglophone Africa, that of Burkina to all Sub-Saharan Francophone Africa and the model can also be applied to Sub-Saharan Lusophone Africa where \( L_1 \) is the native language of the population, \( L_{\text{int}} \) is Portuguese/ Spanish, and \( L_3 \) either French or English.

REFERENCES


Kambou: Contrastive analysis in multilingual societies


