

A synchronic lexical study of the Ede language continuum of West Africa: The effects of different similarity judgment criteria¹

Angela Kluge, SIL International

Summary

In the context of a synchronic lexical study of the Ede varieties of West Africa, this paper investigates whether the use of different criteria sets to judge the similarity of lexical features in different language varieties yields matching conclusions regarding the relative relationships and clustering of the investigated varieties and thus compatible recommendations for further sociolinguistic research. Word lists elicited in 28 Ede varieties were analyzed with the inspection method. To explore the effects of different similarity judgment criteria, two different similarity judgment criteria sets were applied to the elicited data to identify similar lexical items. The quantification of these similarity decisions lead to the computation of two similarity matrices which were subsequently analyzed by means of correlation analysis and multidimensional scaling. The analysis findings suggest compatible conclusions regarding the relative relationships and clustering of the investigated Ede varieties. However, the matching clustering results do not necessarily lead to the same recommendations for more in-depth sociolinguistic research, when interpreted in terms of an absolute lexical similarity threshold. The indicated ambiguities suggest the usefulness of focusing on the relative, rather than absolute in establishing recommendations for further sociolinguistic research.

Zusammenfassung

Im Rahmen eines synchronen lexikalischen Forschungsprojekts der Varianten des Ede in Westafrika, untersucht der Aufsatz, ob die Anwendung unterschiedlicher Sets von Kriterien zu entsprechenden Ergebnissen in Bezug auf die verwandtschaftlichen Beziehungen und Gruppierung der betroffenen Varianten liefert. Somit erhebt sich die Frage, ob der Ansatz für zukünftige soziolinguistische Untersuchungen empfehlenswert ist. Wortlisten aus 28 Ede-Sprachen wurden mit der Inspektionsmethode untersucht. Um die Auswirkungen verschiedener Kriterien für die Ähnlichkeitsprüfung auszuwerten, wurden zwei unterschiedliche Sets von Kriterien zur Ähnlichkeitsprüfung auf die erhobenen Daten für die Bestimmung ähnlicher Wörter angewendet. Die Quantifizierung dieser Ähnlichkeitsentscheidungen führte zur Berechnung von zwei Ähnlichkeitsmatrizes die dann mittels einer Korrelationsanalyse und einer mehrdimensionalen Skalierung analysiert wurden. Die Ergebnisse dieser Untersuchung lassen vergleichbare Schlüsse in Bezug auf die Verwandtschaftsbeziehungen und Gruppierungen der untersuchten Varianten des Ede erwarten. Doch selbst wenn die Gruppierungen passen, bedeutet dies nicht unbedingt, dass dieselben Empfehlungen auch für mehr in die Tiefe gehende soziolinguistische Untersuchungen gelten. Die aufgezeigten Zweideutigkeiten lassen es sinnvoll erscheinen, sich bei der Erarbeitung von Empfehlungen für weitere soziolinguistische Untersuchungen mehr auf das Relative als das Absolute zu konzentrieren.

Résumé

Dans le cadre d'un projet de recherche de lexique synchronisé des variantes d'Ede en Afrique de l'Ouest, cette étude essaie de voir si l'emploi de différents lots de critères donne des résultats appropriés concernant les relations de parenté et le regroupement des variantes en question. Ainsi, on se pose la question, à savoir s'il est recommandé de faire une tentative de recherches sociolinguistiques futures sur le sujet. Des listes de mots établies à partir de 28 langues Ede ont été examinées avec la méthode d'inspection. Pour explorer les effets de différents lots de critères dans l'examen de la ressemblance, on a utilisé deux différents critères d'examen en vue de déterminer les mots qui se ressemblent. La quantification de ces décisions de similitude a amené à concevoir deux matrices de ressemblance, qui seront ensuite examinées à l'aide d'une analyse de co-relation et d'une graduation à plusieurs dimensions. Les résultats de cette investigation laissent espérer des conclusions similaires en ce qui concerne les relations de parenté et le regroupement des variantes d'Ede qui ont été examinées. Cependant, la réussite des regroupements ne signifie pas qu'il y aura obligatoirement les mêmes recommandations pour des recherches sociolinguistiques plus poussées. Les ambiguïtés indiquées suggèrent qu'il serait mieux, dans la recherche de recommandations pour d'autres investigations sociolinguistiques, de se concentrer beaucoup plus sur le relatif que sur l'absolu.

Introduction

<1>

This paper describes the findings of a synchronic lexical study of the Ede varieties of West Africa. In presenting the findings, this paper explores the clustering of the Ede speech varieties and discusses the role of different sets of similarity judgment criteria in analyzing lexical features. In the context of this analysis, special attention is given to the issue of complex polymorphemic word-structures. More specifically, the question is explored whether different sets of similarity judgment criteria lead to compatible conclusions as to the relative relationships and clustering of the investigated varieties and recommendations for further sociolinguistic research.

<2>

The Ede language continuum (Benue-Congo language family) is situated in the southeastern part of West Africa. Expanding westward from southwestern Nigeria, the Ede language communities occupy large areas of central Benin, as well as two pockets in the eastern part of central Togo. Thus far, two of the Ede varieties have undergone language development on a larger scale: Ifè in Togo, and Yoruba both in Benin and Nigeria. To assess extensibility issues, that is to investigate whether the remaining Ede communities could benefit from these literacy efforts or whether additional development programs in some of the remaining communities would be beneficial, a sociolinguistic study of the language continuum was launched in the early 1990s by the Togo-Benin branch of SIL International.

<3>

During the first phase of this study, which is the focus of this paper, word lists were elicited in 28 Ede varieties of Togo, Benin, and Nigeria. Main objectives of these elicitations were: (1) to obtain a rough estimate of the computed degrees of linguistic similarity between these varieties, (2) to explore how these varieties might be treated as clusters, and (3) to establish recommendations for further sociolinguistic research.

<4>

A synchronic approach was chosen for the analysis of the elicited lexical Ede data, given the study's overall objective of assessing the extensibility of the literacy efforts already existing. The main tool for this approach is the inspection method, as first described by Gudschinsky (1955). This method focuses on phonetic similarity and measures the relative degrees of lexical relationships of closely related languages which are in turn taken as indicators of potential intelligibility. Gudschinsky (1956: 206) concludes that knowledge of the degree of lexical similarity between speech varieties "*is invaluable in practical decisions regarding homogeneity of speech areas for vernacular schools, production of literature, etc.*" (See also Sanders 1977, Saussure 1959, and Simons 1977).

<5>

It cannot be assumed, though, that high degrees of lexical similarity correlate with high degrees of intelligibility. However, as pointed out by Grimes (1988) in his study on 'Correlations between vocabulary similarity and intelligibility,' the opposite relationship does hold: a low degree of lexical similarity always correlates with a low degree of intelligibility. Grimes (1988) concludes further that while high degrees of lexical similarity do not provide valid, reliable indicators for high degrees of intelligibility, lexical similarity degrees can point to those areas where more in-depth studies are necessary.

<6>

Therefore, for further interpretation of word list results in the context of extensibility studies, SIL's "Language assessment criteria" (International Language Assessment Conference 1990) recommend that if lexical similarity is more than about 70% (at the upper confidence limit of the calculation), dialect intelligibility testing should be conducted to assess intelligibility degrees between the speech varieties under investigation, whereas lexical similarity of less than 70% would indicate that the varieties in question are sufficiently different to be regarded as distinct languages.

<7>

With respect to the procedures involved in making lexical similarity decisions, Gudschinsky (1956) provides a set of guidelines that require a thorough morphemic analysis of the elicited lexical items to isolate equivalent morphemes and subsequently to identify phonetically and semantically similar lexical items and to group these into sets of probable cognates. These guidelines have since been

adapted for use in linguistic and sociolinguistic language surveys, for example in Blair's (1990:30ff) manual for small-scale language surveys which provides a set of criteria for the comparison of pairs of phones in two words. In contrast to Gudschinsky's (1956) guidelines, though, Blair's (1990) guidelines, while taking into account word length, do not deal with the morphemic structure of the elicited lexical items, for example how to handle complex polymorphemic word-structures such as additional morphemes and reduplication.

<8>

Ideally though, a thorough morphemic analysis that isolates the equivalent basic stem morphemes in each pair of words should form the basis for any lexical similarity decisions. However, in the context of most sociolinguistic language surveys that are limited in scope and time and focus on the extensibility of potential or already existing literacy efforts, such thorough morphemic analysis proves unfeasible in most cases, especially for languages with complex polymorphemic word-structures (Probst 1992, and Sanders 1977).

<9>

Thus, to a large extent SIL International language survey teams have been employing the inspection method in their lexical analyses, applying the principles outlined by Blair (1990) to make similarity decisions without thorough morphemic analysis (see for example Dettweiler & Dettweiler 2003, Hochstetler et al. 2004, Meyers et al. 2003, Rueck and Jore 2003). The analysis findings are subsequently interpreted in terms of SIL's above-mentioned, recommended 70%-threshold level to establish recommendations for further sociolinguistic research: if lexical similarity is lower than 70%, it is assumed that the speech communities in question would not understand the other speech forms well enough to be able to benefit from the same efforts, whereas if lexical similarity is higher than 70%, dialect intelligibility testing is recommended to assess whether speakers can understand the other speech forms well enough to be able to benefit from the same literacy efforts (see for example Brye & Brye 2004, Dettweiler & Dettweiler 2003, Harrison et al. 1999, and Hochstetler et al. 2004).

<10>

The fact though, that the established analysis procedures for lexical similarity decisions take into account word length but not the morphemic structure of the elicited items does raise the question how to handle the lexical analysis of languages with complex polymorphemic word-structure. Moreover, the threshold of 70% lexical similarity brings up the question whether and to what extent the application of different similarity judgment criteria in analyzing these lexical items would yield differing conclusions regarding the relative relationships and clustering of the investigated varieties which in turn would result in differing recommendations for further sociolinguistic research.

<11>

These questions were explored in a previous study on the Gbe language continuum of West Africa (Kluge 2005). Word lists were elicited in 49 Gbe varieties and analyzed by means of the inspection method. To explore the effects of different similarity judgment criteria, two different similarity judgment criteria sets were applied to the elicited data to identify similar lexical items. In a next step, these similarity decisions were quantified and two similarity matrices computed. These matrices were subsequently analyzed by means of correlation analysis and multidimensional scaling. The findings of the correlation analysis indicated a significant linear and positive relationship between both word list computations. These findings supported the conclusion that application of either set of similarity judgment criteria would lead to similar clustering results for the Gbe data set. This conclusion was substantiated by the findings of multidimensional scaling which suggested that different sets of similarity judgment criteria lead to similar clustering results and similar recommendations for further research.

<12>

This previous study, dealing with but a single set of speech varieties, was presented as a case study for investigating the effects of different criteria sets and the usefulness of focusing on the relative relationships and clustering of the investigated varieties in establishing recommendations for more in-depth sociolinguistic research. Kluge (2005) concluded that comparable studies in different language situations, applying the techniques suggested for the Gbe data set, were needed to verify the validity of this data analysis approach.

<13>

Thus, the main objectives of the current analysis of the Ede data set are twofold: (1) to assess the clustering of the Ede speech varieties, and (2) to explore the effects of different similarity judgment criteria and thereby verify the validity of the data analysis approach previously applied to the Gbe data. Therefore, two different sets of similarity judgment criteria were applied to the Ede data set to explore their effects on the computed degrees of lexical similarity, initial clustering of these varieties, and recommendations for further sociolinguistic research.

<14>

After presenting, in Section 2, pertinent background information on the Ede language continuum, Section 3 describes the methodology employed in analyzing the elicited lexical features which resulted in the computation of two lexical similarity matrices based on the two different sets of similarity judgment criteria. Section 4 presents the findings of this analysis, focusing on the statistical relationships between both computations and the clustering of the Ede varieties according to both criteria sets. In Section 5, the findings of this analysis are discussed with special emphasis given to the question whether the application of different similarity judgment criteria sets leads to the same or different conclusions.

Background information on Ede

Language classification

Capo (1989:281), building on Akinkugbe's (1978) and Williamson's (1989) work, and based on personal field notes, proposes the following classification for the Yoruboid language varieties:[^]

Niger-Congo, Atlantic-Congo, Volta-Congo, (New) Benue-Congo, Defoid, Yoruboid

1. Edekiri (Yoruba/Isekiri)
 - a) Ede (Yoruba)
 - b) Isekiri / South East Ede
2. Igala

<15>

Within the Yoruboid language group, and especially the Edekiri branch, several not-well defined terms are in use. Capo (1989:277) mentions that the term 'Yoruba' can be used as:

<16>

... a cover term for a dialect cluster spoken in Western Nigeria, Bénin and Togo. However, in actual fact it seems not to cover the speakers of the cluster in Bénin and Togo because they are better known as Nagó or Anagó. For others, Yorùbá is a separate member of the dialect cluster actually referring to the written form, the standard variety accepted by most of the dialect speakers.

<17>

Given these sometimes ambiguous glossonyms and to avoid the term 'Yoruba' when referring to the Yoruboid branch, Capo (1989:281) proposes 'Ede' as a cover term for the Yoruba dialect cluster, since all varieties of this cluster prefix 'èdè', the standard Yoruba term for 'language', to their respective ethnonyms (Capo 1989: 281), a proposal that was "*tacitly accepted at the 8th LAN Conference at Port Harcourt*" in August 1987. Following this proposal, the term 'Ede' rather than 'Edekiri' or 'Yoruba' will be employed throughout this paper when referring to the Yoruba dialect cluster, while, following Capo's definition (1989:282), the term 'Yoruba' will be employed when referring to "*the standard language of the whole area*".

<18>

In his overview of the Yoruboid language family, Capo (1989:279-281) provides a preliminary listing of Ede speech varieties which are presented in Table 1, according to the country where they are spoken (see also Figure 6 for a map of the Ede language area, Section 5.1):²

Table 1: Listing of Ede (Yoruba) varieties according to Capo (1989:279-281)

Benin	Nigeria		Togo
Ajashe	Awori	Ijumu	Ana (Ifè)
Ana (Ifè)	Bunu	Ikale	Ica
Cabe	Egba	Ilaje	
Ica	Egbe	Oba	
Idaca	Ekiti	Oka	
Ije	Gbede	Ondo	
Ketu	Ibolo	Oshun	
Mokole	Ifaki	Owe	
Nago	Ifè	Owo	
	Igbomina	Oworo	
	Ije	Oyo	
	Ijebu	Ukare	
	Ijesha	Yagba	

<19>

Capo (1989:281) further proposes the internal classification of the Ede language varieties into four subgroupings. He comments though that this classification is presented with some reservations since, in his view, “*not enough information is available to arrive at reliable subclassifications*” on the basis of “*rigorous sound laws and specific innovations*”; instead, “*linguistic evidence is found to support existing geographical and ethnopolitical subgroupings*” (Capo 1989:283):

1. Central Ede: Ifè, Ijesha, Ekiti, etc.
2. Northeast Ede: Yagba, Gbedde, Ijumu, etc.
3. Southwest Ede: Cabe, Ketu, Ana, etc.
4. Northwest Ede: Oyo, Egba, Oshun, etc.

<20>

Other listings of Ede varieties are provided by the Ethnologue (Gordon 2005) and, for Benin, by the Carte linguistique, a language map designed by Benin’s national center for applied linguistics (Centre National de Linguistique Appliquée, CENALA 1990).

<21>

The Ethnologue (Gordon 2005) proposes the internal classification of the Edekiri speech varieties into five groupings: Ede, Ifè, Isekiri, Mokole, Ulukwumi, and Yoruba (see Table 2).³

Table 2: Listing of Edekiri varieties and their subgroupings according to the Ethnologue (Gordon 2005)

Benin	Nigeria	Togo
Ede	Isekiri [its]	Ede
Cabe [cbj] ¹	Ulukwumi [ulb]	Nago, Manigri-Kambolé [xkb]
Ica [ica]	Yoruba [yor]	Ifè [ife]

¹ [...] = ISO 639-3 code (Gordon 2005).

Benin	Nigeria	Togo
Idaca [idd]	Akono Ikale	
Ije [ijj]	Awori Ila	
Nago [nqg]	Aworo Ilaje	
Nago, Kura [nqq]	Bunu Iworro	
Nago, Manigri-Kambolé [xkb]	Egba Jumu	
Ifè [ife]	Ekiti Ondo	
Mokole [mkl]	Gbedde Owe	
Yoruba [yor]	Igbonna Oyo	
Egba	Ijebu Wo	
	Ijesha Yagba	

<22>

CENALA's (1990) Carte linguistique indicates the following Ede language areas of Benin: Cabe, Ica, Idaca, Ifè, Ije, Mokole, Nago, and Yoruba.

<23>

Disparities between the Ethnologue (Gordon 2005), the Carte linguistique's (CENALA 1990) and Capo's (1989) listing of Ede varieties refer to Ajashe, Ifè, Mokole, Nago, and Yoruba as well as to Isekiri and Ulukwumi.

1. Ajashe

Capo (1989) lists Ajashe as an Ede variety of southeastern Benin spoken in the town of Porto-Novo, whereas CENALA's (1990) Carte linguistique designates Ajashe as Yoruba; the Ethnologue (Gordon 2005) makes no mention of Ajashe but lists Yoruba with its Egba dialect as the Edekiri language spoken in Porto-Novo. (See also #0 'Yoruba' below.)

2. Ife

Capo (1989:279) gives 'Ifè' as an alternative name for 'Ana', while CENALA's (1990) Carte linguistique designates Ana as Ifè. The Ethnologue (Gordon 2005) does not classify Ifè as an Ede variety, but lists Ifè directly under the Edekiri node, giving 'Ana' as an alternative name.

3. Mokole

Capo (1989:280) includes Mokole in his listing of Ede varieties, specifying that it is spoken "in the rural district of Kandi, Borgou province, Bénin". Capo (1989:281) does not, however, mention Mokole in his internal classification of the Ede language varieties. In contrast, the Ethnologue (Gordon 2005) does not classify Mokole as an Ede variety, but lists Mokole directly under the Edekiri node.

4. Nago

The Ethnologue (Gordon 2005) lists three distinct Ede Nago varieties: (1) Nago, spoken in villages and towns spread across the *sous-préfectures* of Adja-Ouèrè, Ifangni, Ikpilè, Kétou, Pobè, and Sakété (Ouémé province in southeastern Benin); (2) Manigri-Kambolé Nago, spoken in and around the towns of Manigri (Atakora province in northwestern Benin) and – just across the Benin-Togo border – Kambolé (Centre region in northeastern Togo); and (3) Kura Nago, spoken in and around the towns of Alédjo-Koura, Pélébina, and Sèmèrè (Atakora province in northwestern Benin).

Capo (1989) and CENALA's (1990) Carte linguistique also indicate Nago as a distinct Ede variety, without suggesting a further subclassification though; instead, here Nago subsumes the three distinct Nago varieties mentioned above. In addition, Capo (1989) lists Kétou as a distinct Ede variety, referring to the Kétou *sous-préfecture* in Benin's Ouémé province, whereas the Ethnologue (Gordon 2005) lists Kétou as one of the main Nago speaking towns

but not as a distinct Ede variety.

5. Yoruba

Capo (1989) mentions 26 distinct Ede varieties of Nigeria. In contrast, the Ethnologue (Gordon 2005) does not list any Nigerian Ede varieties, but does mention three distinct Edekiri varieties, that is Isekiri, Ulukwumi, and Yoruba. Under the Yoruba entry, the Ethnologue (Gordon 2005) provides a listing of dialects which concurs to a large extent with Capo's (1989) listing of distinct Nigerian Ede varieties: Awori, Bunu (Bini), Egba, Ekiti, Gbedde, Ijebu, Ijesha, Ikale, Ilaje, Ondo, Owe, Oyo, and Yagba. In addition, the Ethnologue (Gordon 2005) lists Akono, Aworo, Igbonna, Iworro, Jumu, and Wo as dialects of Yoruba, neither of which are mentioned by Capo (1989), but the Iworro, Jumu, and Wo varieties could potentially refer to Capo's (1989) Ede varieties Oworo, Ijumu, and Owo, respectively. Capo (1989) also lists a number of Nigerian Ede varieties that are not mentioned by the Ethnologue (Gordon 2005), neither as dialects of Yoruba nor as distinct Edekiri or Ede entries: Egbe, Ibolu, Ifaki, Ifè, Igbomina, Oba, Oka, Oshun, Ukare.

For Benin, the Ethnologue (Gordon 2005) lists a separate entry for Yoruba, referring specifically to the Egba dialect, which is mainly spoken in Porto-Novo and "*throughout the country in the towns and major villages*". (See also #0 '1.' above.)

Isekiri and Ulukwumi

The Ethnologue (Gordon 2005) lists Isekiri and Ulukwumi as two distinct entries directly under the Edekiri node. In contrast, Capo (1989) classifies Isekiri as a variety of the 'Isekiri / South East Ede' grouping which is distinct from the Ede (Yoruba) grouping; Ulukwumi is listed as a Defoid variety but not further classified.

Language area

<24>

The Ede⁴ people originate from southwestern Nigeria from where they migrated westward into today's Benin and Togo. According to tradition, as described by Parrinder (1947) in Yoruba-speaking peoples in Dahomey, this migration which took place hundreds, perhaps even thousands of years ago, originated in the region of Ilesha (Oyo State, Nigeria). However, there have been other waves of immigrants from Nigeria during the last century or two, this time Egba people from the region of Abeokuta (Ogun State). From the town of Porto-Novo (Benin), located near the Nigerian border, the immigrants spread northward. Taking in Sakété, Pobè and Kétou, they spread right across the ancient kingdom of Abomey until they reached the towns of Dassa-Zoumé, Savè, and Kilibo, extending as far north as Savalou. Still further north there are even scattered Ede groups in the area of Djougou and Bassila, which, neighboring the Kotokoli language area, marks an approximate limit. Descendants of these first immigrants referred to themselves as "*Sha*" or "*Itsha*" which might still be seen from the names of some of their principal towns, such as Da-sha (Dassa), Sha-bè (Savè), Sha-bè-lu (Savalou), or Ba-shida (Bassila).

<25>

From here, the immigrants spread further westward into the Ewe language area as far west as Atakpamé, in today's Togo. Other immigrant groups who had reached Bassila in western Benin spread back eastward as far as the Bariba language area with the towns of Alafia and Tchaourou, both names being Yoruba formations, marking an approximate limit. (See Parrinder 1947.)

<26>

Thus, today's Ede speech communities are located in three countries in the southern part of West Africa expanding westwards from southwestern Nigeria across Benin to the eastern part of central Togo (Capo 1989, Gordon 2005).

<27>

The Ede language communities of Nigeria occupy the southwestern part of the country, namely most of the Kwara, Lagos, Ogun, Ondo Osun, and Oyo states, and the western local government area of Kogi State (Gordon 2005). From here the language continuum expands westwards across the border

into the Benin's Ouémé province where the Ije and southern Nago communities are located. These communities are separated from Benin's remaining Ede varieties by Ouémé river. West of the Ouémé river, in the Zou province, the Cabe and Idaca communities are located. Still further west, on both sides of the Benin-Togo border, another group of Ede language communities is located: the Ica, Ifè, and northern Nago communities (i.e. Manigri-Kambolé Nago). North of these communities, in the Atakora province of northwestern Benin, another pocket of Ede varieties is located, the Kura speech communities. Still further north, ca. 280 km north of the Kura and Cabe communities, in Benin's Borgou province, is the Mokole language area located. (See Capo 1989 and Gordon 2005.)

<28>

The Ede communities of Nigeria have as neighbors other Benue-Congo language groups to the east and north and Mande language groups to the northwest. In Benin and Togo, the Ede communities are neighbored by Kwa language groups to the south and Gur and Kwa language groups to the north (Capo 1989: 278, Gordon 2005).

Population

<29>

For Nigeria, the *Ethnologue* (Gordon 2005) gives a total of 18,850,000 Yoruba speakers. The population total for all countries is estimated with 19,327,000 Yoruba speakers and the total number of second-language speakers of Yoruba is estimated with an additional 2,000,000 speakers.

<30>

In Benin, during the 1992 Census, population data were elicited giving totals both by ethnic group, as well as by political community. The total by ethnic group numbers the population of the category 'Yoruba et Apparenté' (Yoruba and Related) as 594,776 (287,767 males and 307,009 females) (Ministère du Plan 1994:47). It is noted though that during the census individuals were not asked which language they speak as their first language but to which ethnic group they belong. Thus, interviewees identified with their father's ethnic group, even though they might not speak his language nor live in his language area. The *Ethnologue* (Gordon 2005) provides a higher estimate of Ede speakers of Benin, giving a total of 1,147,000 speakers for all listed Ede varieties (69,000 Cabe, 63,000 Ica, 100,000 Idaca, 80,000 Ifè, 50,000 Ije, 65,000 Mokole, 200,000 Nago, 25,000 Nago Kura, 30,000 Nago Manigri-Kambolé, and 465,000 Yoruba).

<31>

For Togo, the *Ethnologue* (Gordon 2005) gives a total of 142,000 Ede speakers, including 102,000 Ifè speakers and 40,000 Nago Manigri-Kambolé speakers.

Previous linguistic research

<32>

As regards the Defoid language group as a whole, an inventory and a classification of all the languages of the branch, as well as a general summary of the phonological characteristics of the Defoid language group are given by Capo (1989) (see Section 2.1).

<33>

In a more recent publication, Baloubi (2004) briefly describes the Yoruboid speech varieties of Benin, separating them into four major groups: Benin Central Yoruboid, Benin Northern Yoruboid, Benin Southeastern Yoruboid, and Nigeria Southwestern Yoruboid. Benin Central and Northern Yoruboid (BCNY) refer to the Ica, Idaca, Ifè, Manigri, Cabe, and Mokole varieties, all of which, with the exception of Mokole, "are geographically an extension of the BSY" (Baloubi 2004:52). These varieties are close to the Nigeria Southwestern Yoruboid (NSY) varieties, both in terms of their linguistic features as well as their geographic distance. Benin Southeastern Yoruboid is comprised of the Yoruboid speech varieties of southeastern Benin and refers to the Yoruboid communities of Kétou, Pobè, Porto-Novo, and Sakété.

<34>

As far as linguistic descriptions of individual Ede languages are concerned, overall, little research has been done outside the realm of standard Yoruba. Focusing on Benin and Togo, specific Ede varieties that have received attention are Cabe, Ica, Idaca, and Ifè.

<35>

Regarding the Cabe and Ica varieties, some linguistic analysis was done on Cabe in the context of a comparative phonology of Yoruba dialects, Isekiri, and Igala (Akinkugbe 1978) and on Ica in the context of an MA thesis that explored the meaning of personal nouns (Aguidi 1992).

<36>

Concerning Idaca, reference must be made to the work by Odoun Kouyomou, whose work includes a description of the Idaca phonology (1986) and of various aspects of the grammar (1991c), preliminary sketches of an Idaca-French lexicon (1990, 1991d,e) as well as an introduction to literacy development (1989,1991b). More recently, Baloubi (2004) presented some work on Idaca, exploring the linguistic and ethnic identity of these communities.

<37>

Regarding Ifè, linguistic analysis has focused on various aspects of the phonological system (Boëthius 1983, Kohler 1983), the grammar (Boëthius 1987, Klaver 1995, 1999), and discourse (Klaver 1987). In addition, descriptions and discussions of various aspects of Ifè literacy are available such as by Boëthius (1991, 1987) and Odoun Kouyomou (1991a). (For a more complete listing see SIL International 2007).

<38>

In addition to these linguistic descriptions, word lists were elicited in various Ede varieties of Benin, such as Cabe, Ica, Idaca, Ifè, Ije, Mokole, Nago from Kétou, and the urban variety of Yoruba (Comision Nationale de Linguistique 1983).

<39>

Besides linguistic research, different aspects of the culture and history of a number of Ede varieties of Benin have been studied, among which are Cabe, Ica, Idaca, Ifè, Ije, Mokole, Nago, and Yoruba of Benin. (For more details, see Kluge 1999, Appendix 5.)

Language development

<40>

Thus far, two Ede language varieties have been targeted for language-based development on a relatively large scale, Yoruba in Nigeria and Benin and Ifè in Togo. In Nigeria, Yoruba is the official language for the southwestern part of the country with standard Yoruba being used for government notices, mass media such as newspapers, broadcasting and television, and primary and secondary school education. (Gordon 2005)

<41>

In Benin and Togo, where French is the official language, Yoruba and Ifè, respectively, have been targeted mainly for non-formal adult education.

<42>

In Benin, the Yoruba language, together with Aja, Bariba, Dendi, Ditammari, and Fon, was selected in 1992 by the government for large-scale non-formal adult education (Direction de l'Alphabétisation 1992). Given that Yoruba is linguistically relatively close to Benin's remaining Ede varieties and serves as the language of wider of communication for these speech communities, non-formal adult education programs among the remaining Ede speech communities are also conducted in Yoruba, presupposing that Yoruba literature would appropriately extend to use by these remaining Ede speech communities.

<43>

In Togo, government sponsored non-formal education has previously focused on French, Ewe, Kabiye, Moba, and Tem. However, in the early 1980s, Ifè language development was started in Togo by SIL Togo-Benin, with the Tchetti variety, located in Benin close to the Benin-Togo border, chosen as the reference dialect. Following an initial phase during which several aspects of the language were

analyzed and an alphabet was established with the collaboration of a local language committee, an Ifè literacy program was begun in 1989. By 2004, the program had graduated well over 4,000 Ifè men and women with interest in reading growing steadily. At that time, 45 books in Ifè were available consisting of a core of didactic materials as well as a selection of general reading materials. (Devine 1996, 2004, Klaver 1997)

Methodology

<44>

Based on Capo's (1989) work on the larger Ede language continuum, as well as based on Boëthius' (1983) and Kohler's (1983) research on Ifè in Togo, 163-word lists were elicited in 28 Ede varieties. The word list used was based on Swadesh's 100-word list (Swadesh 1955) and the word list published in the *Atlas linguistique du Cameroun* (Dieu and Renaud 1983). The analysis of the elicited 28 Ede word lists was conducted in four steps. The first step consisted of a qualitative analysis and focused on the identification of similar lexical items. Steps two to four include the quantitative data analysis consisting of the computation of two similarity matrices, a paired t-test and correlation analysis, and multidimensional scaling of both matrices. In addition, a descriptive analysis was conducted to compare the findings of multidimensional scaling to the computed lexical similarity percentage matrices.

Identification of similar lexical items

<45>

During this first step of the analysis, the elicited word lists were analyzed from a synchronic perspective employing an inspection method to assess the relative lexical similarity degrees among the Ede varieties under investigation.

<46>

Allowing for a few modifications, the lexical similarity decisions were based on the principles provided by Blair (1990:31ff). Following these guidelines, two lexical items are considered phonetically similar if at least half of the segments compared are the same or very similar⁵ and of the remaining segments at least half are rather similar.⁶ These guidelines do not consider the morphemic structure of lexical items, and therefore the question arose of how to deal with the polymorphemic word-structure for a fair number of the elicited items.

<47>

In a paper reviewing various methods of handling multimorphemic words in lexicostatistics, Probst (1992) draws special attention to two approaches. The first one requires a thorough morphophonemic analysis, which proves rather unfeasible in the context of sociolinguistic language surveys limited in scope and time. The second method, suggested by Schooling (1981), does not require a morphological analysis. In a survey of French Polynesia, Schooling (1981) compared lexical items as a whole and discarded reduplication and additional morphemes that occur in the same position.

<48>

Given that a fair number of the elicited Ede data are marked by reduplication and additional morphemes and also given the limited scope of the lexicostatistical part of the larger Ede study, Schooling's (1981) approach was chosen as the basis for the similarity judgments. Thus, a set of similarity judgment criteria (Criteria Set 1) was established that, following Schooling's (1981) guidelines, disregarded differences in the morphological structure of the elicited items:

1. Additional morphemes were disregarded if they occur in the same position.
2. Reduplication, regardless of the nasalization of the vowel, was disregarded.
3. Class prefixes on nouns were disregarded.

<49>

To explore how different similarity judgment criteria affect the lexicostatistical comparison of polymorphemic words, Probst (1992) proposed that a second, more rigorous criteria set be applied to the same data set. Thus, a second criteria set was established that, following Probst's (1992) suggestions, does not disregard differences in the morphological structure:

1. Pairs of complete words were compared.
2. Additional morphemes were included in the analysis.
3. Reduplication was included in the analysis.
4. Class prefixes on nouns were disregarded.

<50>

Table 3 gives three examples for both similarity judgment criteria sets. The first example considers affixed morphemes, displaying four lexical items for Gloss #141 ‘(to) swim.’ For Criteria Set 1, apparently affixed morphemes occurring in the same position were disregarded. Thus, focusing on the morpheme [wè], the elicited words were considered lexically similar for the Owo, Cabe (from Savè), and Ijebu varieties as well as for the Owo and Ije varieties. However, the items for Ije and Savè-Cabe as well as for Ije and Savè-Cabe were not considered similar since the additional morphemes do not occur in the same position. Applying the second, more rigorous, criteria set to the same data, only the items for Savè-Cabe and Ijebu were considered lexically similar, whereas the items for Owo and Savè-Cabe, the Owo and Ijebu, the Owo and Ije, the Ije and Savè-Cabe, as well as for the Ije and Ijebu varieties were considered lexically non-similar.^

Table 3: Similarity judgment criteria sets

Criteria Set 1				Criteria Set 2			
Affixed morphemes – ‘swim’ (#141)							
Similar		Non-similar		Similar		Non-similar	
Owo:	wè	Ije:	wɛdo	Cabe (S):	mũwè	Owo:	wè
Cabe (S):	mũwè	Cabe (S):	mũwè	Ijebu:	lùwè	Cabe (S):	mũwè
Ijebu:	lùwè	Ije:	wɛdo			Owo:	wè
Owo:	wè	Ijebu:	lùwè			Ijebu:	lùwè
Ije:	wɛdo					Owo:	wè
						Ije:	wɛdo
						Ije:	wɛdo
						Cabe (S):	mũwè
						Ije:	wɛdo
						Ijebu:	lùwè
Reduplication – ‘bark’ (#75)							
Similar		Non-similar		Similar		Non-similar	
Egba:	èkpo			Egba:	èkpo	Idaca:	ìkpoḱp o
Ondo:	ìkpo			Ondo:	ìkpo	Egba:	èkpo
Idaca:	ìkpoḱp o					Idaca:	ìkpoḱp o
						Ondo:	ìkpo
Class prefixes – ‘mouth’ (#1)							
Similar		Non-similar		Similar		Non-similar	
Ica:	anũ			Ica:	anũ		
Ije:	anu			Ije:	anu		

Egba:	enũ			Egba:	enũ		
-------	-----	--	--	-------	-----	--	--

<51>

The second example considers reduplication. For Criteria Set 1, reduplication was disregarded, and thus for gloss #75 ‘bark’ the lexical items were considered similar for the Egba, Ondo, and Idaca varieties. In contrast, for Criteria Set 2, only the items for Egba and Ondo were considered lexically similar, whereas for the Idaca and Egba varieties as well as for the Idaca and Ondo varieties they were considered lexically non-similar.

<52>

Finally, in the third example class prefixes were disregarded for both criteria sets, and thus the elicited items for gloss #1 ‘mouth’ were considered similar for the Ica, Ije, and Egba varieties.

<53>

Following these examples, the complete Ede data set was analyzed with both similarity judgment criteria sets being applied to the elicited data. For 110 glosses (67%) of the 163-item word list, the different treatment of reduplication and additional morphemes resulted in different similarity judgment decisions in at least one language pair.

<54>

Employing WordSurv (Wimbish 1989), a computer program designed for analyzing language survey word lists, the elicited data were organized in two different databases referring to the two different criteria sets.

Computation of WordSurv similarity matrices

<55>

During this second step of the analysis, the similarity judgments established during the first analysis step were quantified. Two lexical similarity percentage matrices were computed each of which is based on a count of shared similar lexical items between each pair of Ede varieties: word list computation 1 (‘WLC1’) is based on Criteria Set 1 and word list computation 2 (‘WLC2’) is based on Criteria Set 2.

<56>

Besides the lexical similarity percentage matrix, WordSurv also computes a variance matrix with range of error for each count. This calculation is based on the reliability of the word list data which takes into account the researcher’s familiarity with the speech varieties under study, availability of good bilingual informants, and opportunities to double-check elicited items (Wimbish 1989:31).

<57>

Given SIL’s language assessment criteria for further interpretation of word list results (see Section 1), the lexical similarity percentages reported in this study refer the sum of the actual measured degree of similarity plus the upper range of error, rather than the actual measured degree of similarity.⁷

Paired t-test and correlation analysis

<58>

During the third step of the analysis a paired t-test was conducted to explore whether the different similarity judgment decisions in the 110 items (glosses) of the 163-item word list yielded significantly different lexical similarity percentages. A major requirement of a t-test is that the observations be independent. Any pair of lexical similarity percentages, however, that have any variety in common are not independent; even pairs that have no variety in common, but which are based on very closely related varieties are not independent. Therefore, to obtain a set of approximately independent observations, 17 location pairs were chosen such that no location is in more than one pair, with each pair consisting of varieties that are about four “steps” away from each other in the matrix of similarity percentages.

<59>

Subsequently, a correlation analysis was conducted. The main objective of this analysis was to examine whether the relationship between the two word list computations is statistically significant, or

in other words whether the two word list computations indicate compatible relative relationships between the Ede speech varieties compared. These analyses were carried out, applying the statistical computer package SPSS for Windows, release 15.0 (SPSS Inc. 2006; see also Norušis 1993, Savage 1999, and Simons 1979) and the package R, release 1.9.1 (R-Project, n.d.).

<60>

To test whether the computed correlation is statistically significantly greater than 0.70, an approximate 95% confidence interval was constructed using the bootstrap method (Efron and Tibshirani 1993). If the lower bound of this interval is greater than 0.70, then the correlation between the two methods is statistically significantly greater than 0.70. Since, as explained above, the full set of observations is not statistically independent, location pairs used for this significance test were 17 selected pairs, the same as chosen for the paired t-test.

Multidimensional scaling

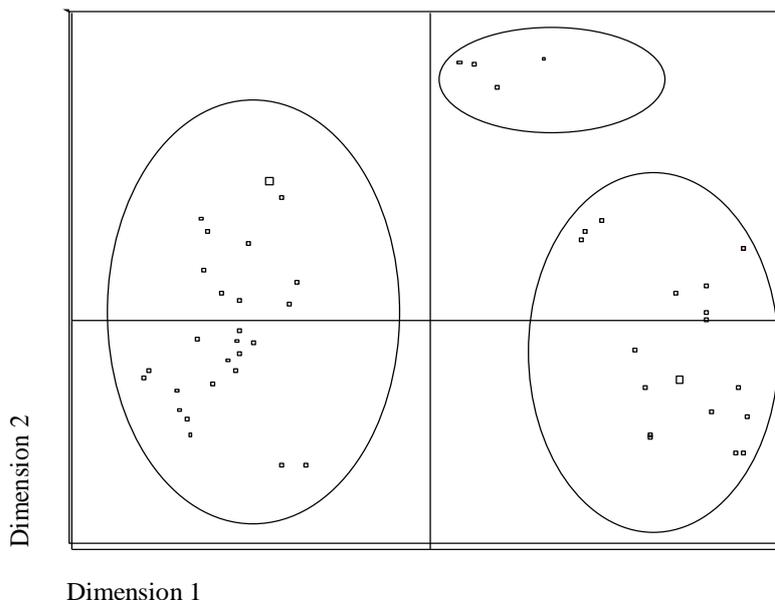
<61>

In the final step of the analysis, the computed word list similarity matrices were analyzed with multidimensional scaling (MDS), again employing SPSS (SPSS Inc. 2006). This statistical procedure is designed to analyze the structure of dissimilarity (or similarity) data with the analysis results presented in a perceptual map, as shown in Figure 1. This perceptual map displays the data as a configuration of points, as on a map, along two, three or more dimensions with the proximity of these points to each other indicating how similar they are.

<62>

The computed perceptual maps do not directly indicate what the dimensions and configurations refer to but it is left to the researcher to interpret what they represent. Most commonly, ‘dimensional interpretation’ and ‘neighborhood interpretation’ are applied to explain as much of the displayed configuration as possible. ‘Dimensional interpretation’ focuses on large distances along the dimensions of the computed plots and requires the researcher to interpret what the dimensions represent. In contrast, ‘neighborhood interpretation’ focuses on data clustering due to large similarities and requires the researcher to identify groups or neighborhoods of stimuli in the multidimensional space (in Figure 1 the identified neighborhoods are indicated by the ovals).

Figure 1: Perceptual MDS map (sample)



<63>

In addition, MDS provides two measures for each configuration to test the results for reliability and validity: (1) the ‘squared simple correlation’ (RSQ) to determine what proportion of variance of the scaled data can be accounted for by the MDS procedure, and (2) a measure of stress (‘Kruskal’s

stress') to measure how well the derived configuration matches the input data. RSQ values of ≥ 0.60 are generally considered acceptable, whereas with Kruskal's measure of stress, small values approaching 0.0 indicate an acceptable goodness of fit.

<64>

According to Woods et al. (1986) and Scholfield (1991), to name but two, MDS can also be applied to the analysis of linguistic data such as speech sounds or words, respectively, in that it provides a measure of distance or dissimilarity for these data by considering all pairs of observations (see Garrett et al. 1999; Hair et al. 1998; Kruskal and Wish 1978; Manly 1986). Thus, for the current study, MDS was employed as an exploratory data analysis tool to investigate the clustering of the Ede varieties based on the computed degrees of lexical similarity as indicated by the computed word list similarity matrices.

<65>

A two-dimensional analysis was applied with the scaling model based on the default 'Euclidean distance measure'. With regard to the interpretation of the computed MDS configurations, a combination of neighborhood interpretation and dimensional interpretation was employed, as suggested by Kruskal and Wish (1978). Thus, the computed perceptual maps were investigated as to what the displayed dimensions represent while at the same time taking into account the clustering of the data.

<66>

Given that MDS was employed as an exploratory data analysis tool, no rigorous validation techniques have been applied. However, an initial descriptive analysis of the computed similarity matrices was conducted to compare the MDS findings to the computed lexical similarity percentage matrices.

Results

<67>

For the current study, 28 Ede varieties were included in the analysis:

Ana (from Sokodé), Boko, Cabe (from Savè and Tchaourou), Egba, Ekiti, Ica, Idaca, Ifè (from Akparè, Atakpamé, and Tchetti), Ije, Ijebu, Ijesham, Ile-Ife, Kura (from Awotébi and Partago), Mokole, Morétan, northern Nago (from Kambolé and Manigri), southern Nago (from Kétou and Pobè), Ondo, Owo, Oyo, Yoruba (from Ibadan and Porto-Novu).

<68>

A qualitative analysis of the 28 elicited word lists resulted in two different word list similarity matrices, WLC1 and WLC2: the computation of WLC1 is based on similarity judgment Criteria Set 1 which disregards differences in the morphological structure, whereas the computation of WLC2 is based on the more rigorous Criteria Set 2 which heeds such differences.

<69>

A first descriptive analysis of the computed similarity matrices indicates a chaining pattern for the entire Ede cluster. The Ede speech groups are situated geographically in a contiguous arrangement from the southwestern corner of Nigeria across southeastern and central Benin into the eastern part of central Togo. Thus, the individual Ede speech groups have contact relationships with the other Ede groups surrounding them which results in the linguistic convergence of adjoining groups: overall, the investigated Ede speech forms are marked by relatively small lexical differences between adjoining dialects whereas differences are greater between Ede varieties at opposing ends of the chain. This chaining pattern is evidenced for both computed word list similarity matrices in which neighboring groups in the chain are more or less placed in adjoining columns and rows with the highest degrees of lexical similarity occurring on the diagonal and the lowest in the corner with degrees of lexical similarity getting sequentially larger from the corner to the diagonal.

Statistical Relationships between WLC1 and WLC2

<70>

Assessing lexical similarity of polymorphemic items, characterized by reduplication and/or additional morphemes, according to the two different similarity judgment criteria sets lead to different similarity

judgment decisions for 110 glosses (67%) of the 163-item word list. For the more rigorous similarity judgment criteria set, WLC2, these different decisions resulted in an overall average degree of lexical similarity of 65% with similarity percentage ranges of 33-95%, whereas for WLC1 the overall average degree of lexical similarity was higher with 75% and the range smaller with 47-98% (n=378), thus yielding an observed difference of 10% between the mean WLC1 percentage and the mean WLC2 percentage.

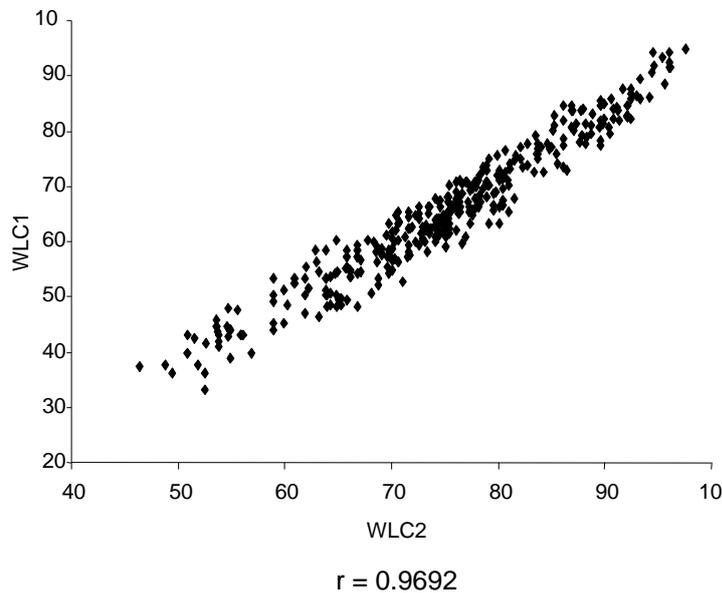
<71>

Results of the paired t-test indicate that the observed mean difference between the WLC1 and WLC2 values for the 17 observations (selected as described in Section 3.3) is 11.04 which is statistically significantly different from 0 (p-value <0.0001). Thus, it can be concluded that the two different similarity judgment criteria sets yield statistically significantly different lexical similarity percentages.

<72>

Correlation analysis of WLC1 and WLC2 results in a narrow scatter, indicating a linear and positive relationship (see Figure 2). This narrow scatter leads to a high correlation coefficient, $r = 0.9692$.

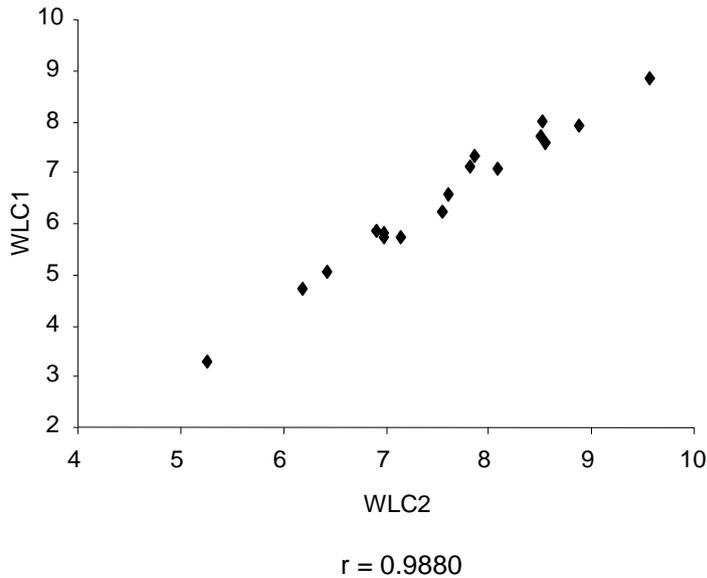
Figure 2: Scatterplot for word list computations WLC1 and WLC2 – Complete data set



<73>

The correlation between WLC1 and WLC2 based on only the 17 observations selected, as described in Section 3.3, is $r = 0.9880$ (see Figure 3). An approximate 95% confidence interval for the correlation between WLC1 and WLC2 is (0.98, 1). Since the lower bound of this confidence interval is larger than 0.70, the correlation between the two similarity measures is statistically significantly greater than 0.70.

Figure 3: Scatterplot for word list computations WLC1 and WLC2 – 17 observations



<74>

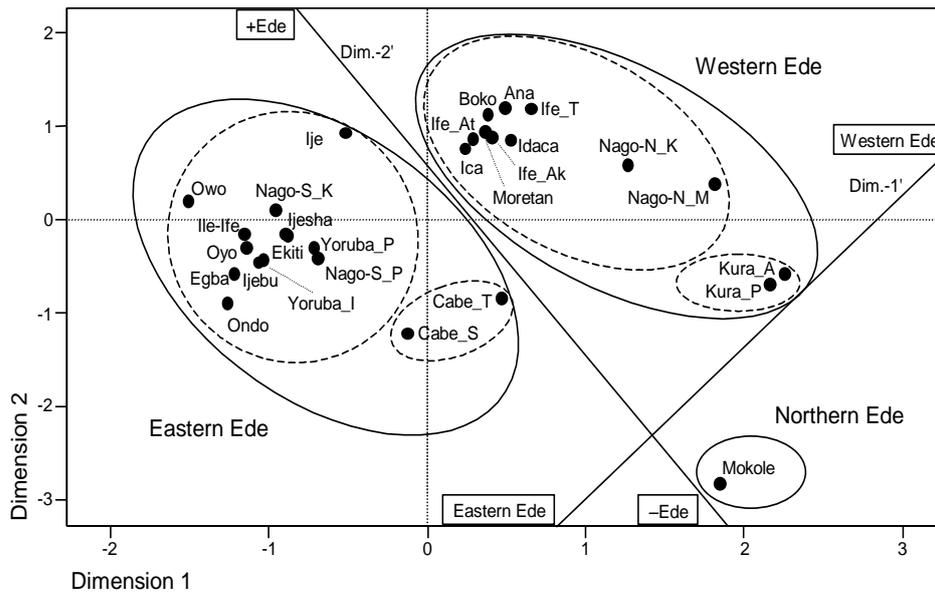
These findings indicate a significant linear relationship between WLC1 and WLC2 with either revealing the same relative distances between the pairs of observed values. Thus, it can be concluded that for the analysis of the Ede data the two different sets of similarity judgment criteria imply the same relative relationships between the investigated speech varieties. These findings support the conclusion that use of either set of similarity judgment criteria would lead to similar clustering results. To verify this conclusion both word list computations were further analyzed with multidimensional scaling.

Clustering of the Ede language varieties

<75>

Multidimensional scaling of the computed word list matrices results in the MDS plots presented in Figure 4 and Figure 5, with the MDS stress values of 0.23975 and 0.26223, and the RSQ values of 0.83694 and 0.79495 for WLC1 and WLC2, respectively, indicating that the derived configurations have an acceptable correlation with the input data. The ovals denote my interpretation of the clusters as indicated by dimensions 1 and 2 of the MDS plots. For both MDS plots, though, dimension 1 and dimension 2 do not lend themselves to a clear interpretation. However, the MDS plot suggests a dimensional interpretation along the superimposed diagonal lines. Thus, according to my interpretation, dimension 2' roughly indicates a degree of "Edeness", whereas dimension 1' suggests the extent to which the Ede varieties are grouped within the Eastern Ede or the Western Ede clusters.

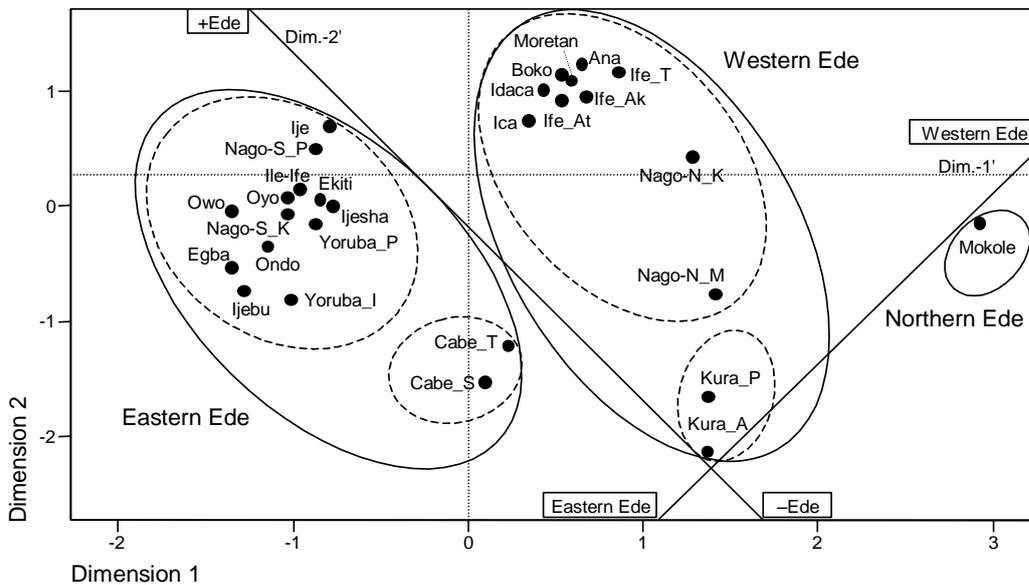
Figure 4: Major Ede clusters – MDS plot of computation WLC1



<76>

For both computations, displayed in Figure 4 and Figure 5, the MDS findings indicate four distinct groupings of Ede varieties.

Figure 5: Major Ede clusters – MDS plot of computation WLC2



<77>

Along the perceptual Eastern-Western Ede axis, the Ede varieties that are situated on the Western Ede side of this axis refer to the Ede varieties west of the Ouémé river in southwestern Benin and southeastern Togo. In contrast, the Ede varieties that are situated towards the Eastern Ede side of the axis refer to the Ede varieties east of the Ouémé river in southeastern Benin and southwestern Nigeria. Thus, the Ouémé river, Benin's largest fluvial basin, serves as a natural barrier between the eastern and western Ede varieties which in the past may have affected contact patterns between the eastern and western Ede varieties, thus resulting in the divergence between the western and eastern Ede speech varieties. Mokole which is situated on the western Ede side of the axis is located in northern Benin, several hundred kilometers north of the remaining Ede varieties. Given their geographical locations, the three Ede clusters are – for the purposes of this paper – tentatively referred to as Western

Ede, Eastern Ede, and Northern Ede. Along the perceptual “Edeness” axis, the eastern and western Ede varieties are situated on the “+Ede” side of this axis, whereas Northern Ede, or Mokole, is placed on its “-Ede” side. Table 4 gives an overview of the identified Ede clusters and their components.

Table 4: Major Ede clusters with elicited Ede varieties according to their larger geographical setting

Western Ede	Eastern Ede	Northern Ede
Located in Benin and Togo west of the Ouémé river	Located in Benin and Nigeria east of the Ouémé river	Located in northern Benin
Ana (from Sokodé)	Cabe (from Savè)	Mokole
Boko	Cabe (from Tchaourou)	
Ica	Egba	
Idaca	Ekiti	
Ifè (from Akparè)	Ije	
Ifè (from Atakpamé)	Ijebu	
Ifè (from Tchetti)	Ijesha	
Morétan	Ile-Ife	
Nago-North (from Kambolé)	Nago-South (from Kétou)	
Nago-North (from Manigri)	Nago-South (from Pobè)	
Kura (from Awotébi)	Ondo	
Kura (from Partago)	Owo	
	Oyo	
	Yoruba (from Ibadan)	
	Yoruba (from Porto-Novo)	

<78>

The placement of the Ede varieties into three major clusters appears to concur with the respective average degrees of intra- and inter-group lexical similarity, indicating the same three major Ede clusters with the same components as the MDS findings.

<79>

Although, due to the different similarity judgment criteria sets lexical similarity percentages differ significantly, a similar pattern emerges for both word list similarity matrices: for both computations, average degrees of intra-group similarity are, overall, higher than average degrees of inter-group similarity, as displayed in Table 5.

Table 5: Average degrees of intra-group and inter-group similarity for the major Ede clusters

Ede cluster	Intra-group similarity ranges		Average intra-group similarity		Ede cluster	Average inter-group similarity	
	WLC1	WLC2	WLC1	WLC2		WLC1	WLC2
West	64-98	48-95	80	71	West – East	72	61
East	66-96	57-94	83	73	West – North	54	44
North	--	--	--	--	East – North	52	40

<80>

For WLC1 and WLC2, average degrees of lexical similarity for the entire Ede cluster are 75% and 65%, with ranges of 47-98% and 33-95%, respectively. Average degrees of inter-group lexical similarity are lower (WLC1: ≤72%, WLC2: ≤61%), whereas average degrees of intra-group lexical similarity for the three identified Ede clusters are higher (WLC1: ≥80%, WLC2: ≥71%). Due to the chaining pattern of the Ede varieties, though, there is some overlap between the largest degrees of inter-group similarity and the smallest degrees of intra-group similarity.

<81>

With regard to intra-group similarity, the MDS findings indicate a somewhat scattered placement of the Western and Eastern Ede varieties, suggesting a fair amount of intra-group variation. These findings are evidenced by the respective ranges of intra-group lexical similarity, as displayed in Table 5, with the range of similarity degrees for the western Ede varieties being slightly larger than for the eastern Ede varieties.

<82>

In the following sections, the MDS findings for both word list computations are discussed in more detail according to their larger geographical clustering, i.e., the western, eastern, and northern Ede varieties.

Western Ede varieties

<83>

The MDS plots situate the western Ede varieties towards the Western Ede side and towards the “+Ede” side of the perceptual Eastern-Western Ede and “Edeness” axes, respectively. The MDS plots also indicate a fairly scattered placement of the western Ede varieties, especially with regard to Kura (from Awotébi and Partago). The Kura varieties are situated the furthest apart from the remaining western Ede varieties and, in comparison to the remaining western varieties, also the furthest down along the “Edeness” axis, towards its “-Ede” side. These distinct placements suggest two subgroupings within Western Ede:

1. Ana, Boko, Ica, Idaca, Ifè (from Akparè, Atakpamé, and Tchetti), Morétan, Nago-North (from Kambolé and Manigri) – due to their geographical location temporarily referred to as southwestern Ede
2. Kura (from Awotébi and Partago) – temporarily referred to as northwestern Ede.

<84>

The grouping of the western Ede varieties into two subgroupings appears to be supported by the respective average degrees of inter- and intra-group lexical similarity for both word list computations, as displayed in Table 6.

Table 6: Average degrees of intra-group and inter-group similarity for the western Ede subgroupings

Western Ede cluster	Intra-group similarity ranges		Average intra-group similarity		Western Ede cluster	Average inter-group similarity	
	WLC1	WLC2	WLC1	WLC2		WLC1	WLC2
Southwest	71-98	62-95	85	79	Southwest – Northwest	67	53
Northwest	--	--	92	82			

<85>

For the entire Western Ede cluster, average degrees of lexical similarity are 80% and 71%, for WLC1 and WLC2, respectively, with fairly large ranges which evidence the fairly scattered placement of the western Ede varieties (WLC1: 64-98%, WLC2: 48-95%).

<86>

With respect to the distinct placements of the southwestern and northwestern Ede varieties, closer investigation of the computed similarity matrices indicates, the average degrees of intra-group lexical similarity are higher both for the southwestern (WLC1: 85%, WLC2: 79%) and the northwestern Ede varieties (WLC1: 92%, WLC2: 82%) than for Western Ede as a whole. At the same time, average degrees of inter-group lexical similarity between the southwestern and northwestern Ede varieties are lower (WLC1: 67%, WLC2: 53%). In consequence, intra-group similarity ranges for southwestern Ede are also smaller (WLC1: 71-98%, WLC2: 62-95%). These findings substantiate the two sub-groupings within the Western Ede cluster.

<87>

The distinct grouping of the Kura varieties might be explained by the location of the Kura communities. Within the Western Ede cluster, the Kura area is located the furthest north. Its closest Ede neighbor is the northern Nago variety of Manigri with which the Kura communities are connected through the International Main Road 3 (RNIE3). This road comes from northwestern Benin and runs southwards, parallel to the Benin-Togo border. It passes through the Kura area and further south through the northern Nago area and still further south through the Ica area. The second closest neighbors to the Kura communities are the Kambolé-Nago and Ana communities, both of which are located in Togo. This northern-most location of the Kura communities may have brought about lower degrees of contact with the remaining western Ede communities which may have resulted in a greater degree of divergence from these varieties which in turn may explain the distinct placement of Kura varieties on the MDS plots. In fact, this placement, together with the rather low degrees of lexical similarity between the Kura varieties and southwestern Ede varieties, renders the grouping of the Kura varieties within the Western Ede cluster somewhat questionable. Moreover, their placement towards the “-Ede” side of perceptual “Edeness” axis raises the question to what extent the Kura varieties are related to the remaining Ede varieties and, in fact, part of the larger Ede cluster. An alternative grouping would be within the Northern Ede cluster together with the Mokole variety. Average degrees of lexical similarity between the Kura varieties and Mokole, however, are low (WLC1: 55%, WLC2: 41%).

<88>

Regarding the southwestern Ede varieties, the MDS plots show a somewhat dense placement of the Ana, Boko, Ica, Idaca, Ifè, and Morétan varieties towards the “+Ede” side of the perceptual “Edeness” axis. At the same time, there is some scatter with regard to the northern Nago varieties (from Kambolé and Manigri). These somewhat scattered placements suggest some inter-group variation between the two northern Nago varieties and the remaining southwestern Ede varieties, a suggestion that appears to concur with the respective computed degrees of lexical similarity, as discussed below.

<89>

For the Ana, Boko, Ica, Idaca, Ifè, and Morétan varieties average degrees of intra-group lexical similarity are 90% (WLC1) and 85% (WLC2). Of these varieties, as already mentioned in Section 2.1, neither Boko nor Morétan are included in Capo’s (1989) or Gordon’s (2005) listings of Ede and Edekiri languages. Ica and Idaca, however, are listed as Ede varieties both by Capo (1989) and by Gordon (2005). In addition, Capo (1989) also lists Ifè as an Ede variety (giving Ifè as an alternative name for the main entry Ana), while Gordon (2005) does not list Ifè as an Ede variety but records it directly under the Edekiri node, giving ‘Ana’ as an alternative name. Given however the high degrees of intra-group lexical similarity for the Ana, Boko, Ica, Idaca, Ifè, and Morétan varieties, the question arises whether the *Ethnologue*’s (Gordon 2005) classification of Ifè as an Edekiri but not as an Ede language is indeed justified.

<90>

For the two northern Nago varieties the degree of intra-group lexical similarity is also rather high 88% (WLC1) and 78% (WLC2). In contrast, average degrees of lexical similarity between the Ana, Boko, Ica, Idaca, Ifè, and Morétan varieties and the northern Nago varieties are lower (WLC1: 77%, WLC2: 68%) with ranges of 71-85% and 62-77% for WLC1 and WLC2, respectively. These ranges are mainly due to the average lexical similarity degrees between Manigri-Nago and the remaining southwestern Ede varieties: among all southwestern Ede varieties, Nago-Manigri has the lowest

average similarity degree with the remaining southwestern Ede varieties (WLC1: 76%, WLC2: 66%). In contrast, all other southwestern Ede varieties have higher average similarity degrees with the remaining varieties of this subgrouping, on average 86% for WLC1 and 80% for WLC2. These findings are evidenced by the MDS plots which situate Manigri-Nago at some distance from northern Nago of Kambolé and the rest of the southwestern Ede varieties towards the “–Ede” side of the “Edeness” axis, when compared to the remaining southwestern Ede varieties. This placement towards the “–Ede” side also implies the placement of Manigri-Nago in somewhat closer neighborhood to the Kura varieties. Closer investigation of the computed similarity matrices indicates, however, lower average degrees of lexical similarity between Manigri-Nago and the Kura varieties (WLC1: 68%, WLC2: 55%) than between Manigri-Nago and the remaining southwestern Ede varieties (WLC1: 76%, WLC2: 66%). Manigri-Nago is also situated in somewhat closer neighborhood to two varieties of the Eastern Ede cluster, that is the Cabe varieties from Tchaourou and from Savè. These placements appear to be substantiated by the computed similarity matrices which show that Manigri-Nago shares its second-highest degree of lexical similarity with Tchaourou-Cabe (WLC1: 82%, WLC2: 73%); Manigri-Nago shares its highest degree of lexical similarity with Kambolé, another northern Nago variety (WLC1: 88%, WLC2: 78%).

<91>

These findings might be explained by the geographical location of the northern Nago communities and the migration history of the Cabe communities. Manigri-Nago is located north of the remaining southwestern Ede communities. This northern-most location may have affected contact patterns with regard to the remaining southwestern Ede communities and, in consequence, may have resulted in a greater degree of divergence from these varieties. Regarding the Cabe migration history, Parrinder (1947) reports that some Ede communities, after having reached Bassila, the sous-préfecture in which the Manigri-Nago communities are located, spread back eastward as far as Tchaourou (see Section 2.2). Having originated from the Manigri-Nago area would therefore explain the somewhat higher degrees of lexical similarity between Tchaourou-Cabe and Manigri-Nago.

<92>

Overall, the somewhat scattered placement of the western Ede varieties indicated by the MDS plots and the degree of variation found in the lexical similarity matrices substantiates the already mentioned chaining pattern for the Ede varieties. The western Ede groups are situated geographically in a contiguous arrangement from southwestern Benin and southeastern Togo northwards with the Ana, Boko, Ica, Idaca, Ifè, and Morétan varieties located in the southern parts of this cluster, the northern Nago varieties north of them, and the Kura varieties still further north.

Eastern Ede varieties

<93>

The MDS plots indicate a fair amount of scatter for the eastern Ede varieties with Cabe (from Savè and Tchaourou) situated the furthest apart from the remaining eastern Ede varieties. These distinct placements suggest two subgroupings within Eastern Ede:

1. Egba, Ekiti, Ije, Ijebu, Ijesha, Ile-Ife, Nago-South (from Kétou and Pobè), Ondo, Owo, Oyo, Yoruba (from Ibadan and Porto-Novo) – due to their geographical location temporarily referred to as southeastern Ede
2. Cabe (from Savè and Tchaourou) – temporarily referred to as northeastern Ede

<94>

The grouping of the eastern Ede varieties into two subgroupings appears to be supported by the respective average degrees of inter- and intra-group lexical similarity for both word list computations, as displayed in Table 7.

Table 7: Average degrees of intra-group and inter-group similarity for the eastern Ede subgroupings

Eastern Ede cluster	Intra-group similarity ranges		Average intra-group similarity		Eastern Ede cluster	Average inter-group similarity	
	WLC1	WLC2	WLC1	WLC2		WLC1	WLC2
Southeast	71-95	62-94	85	77	Southeast – Northeast	76	63
Northeast	--	--	96	91			

<95>

For the entire Eastern Ede cluster, average degrees of lexical similarity are 83% and 73%, for WLC1 and WLC2, respectively, with fairly large ranges which evidence the fairly scattered placement of the eastern Ede varieties (WLC1: 66-96%, WLC2: 57-94%).

<96>

In contrast, average degrees of intra-group lexical similarity are higher both for the southeastern (WLC1: 85%, WLC2: 77%) and the northeastern Ede varieties (WLC1: 96%, WLC2: 91%), whereas average degrees of inter-group lexical similarity between both subgroupings are lower (WLC1: 76%, WLC2: 63%). In consequence, intra-group similarity ranges for southeastern Ede are also slightly smaller (WLC1: 71-95%, WLC2: 62-94%). These findings evidence the two subgroupings within the Eastern Ede cluster.

<97>

With respect to southeastern Ede, the MDS plots show a fairly dense placement of the varieties indicated as components of this cluster, suggesting only little intra-group variation within this grouping. Intra-group similarity ranges for southeastern Ede are not that small though, with 71-95% for WLC1 and 62-94% for WLC2. These somewhat large ranges are mainly due to the computed average degrees of inter-group similarity of 82% (WLC1) and 73% (WLC2) between the Egba, Ekiti, Ijebu, Ijesha, Ile-Ife, Ondo, Owo, Oyo, and Yoruba (from Ibadan and Porto-Novu) varieties on the one side and the Ije and southern Nago (from Kétou and Pobè) varieties on the other side. In contrast, average degrees of intra-group lexical similarity both for the Egba, Ekiti, Ijebu, Ijesha, Ile-Ife, Ondo, Owo, Oyo, and Yoruba varieties (WLC1: 87%, WLC2: 80%) and the Ije and southern Nago varieties (WLC1: 89%, WLC2: 81%) are higher.

<98>

Northeastern Ede refers to the Cabe varieties from Savè and from Tchaourou which share a high degree of lexical similarity (WLC1: 96%, WLC2: 91%). The MDS plots place both varieties at roughly the same distance both from the southeastern Ede varieties and the western Ede varieties. Again, computed lexical similarity degrees support these findings: average degrees of inter-group similarity for Cabe with the southeastern Ede varieties (WLC1: 76%, WLC2: 63%) are the same as with the western Ede cluster (WLC1: 74%, WLC2: 62%). In fact, closer examination of WLC2 indicates that the Cabe varieties share their highest degrees of lexical similarity with two western Ede varieties, that is Ica (Tchaourou-Cabe: 76%, Savè-Cabe: 70%) and northern Nago from Manigri (Tchaourou-Cabe: 73%, Savè-Cabe: 70%). For WLC1, Tchaourou-Cabe again shares its highest degrees of lexical similarity with Ica (86%) and Manigri-Nago (82%), while Savè-Cabe shares equally degrees of high lexical similarity of 81% with Ica and with the southeastern Ede variety Yoruba (from Porto-Novu). Moreover, the Cabe varieties are the eastern Ede varieties situated the furthest down along the “Edeness” axis, towards its “-Ede” side. A closer investigation of the computed similarity matrices appears to substantiate this placement. This investigation indicates that among all eastern Ede varieties, the Cabe varieties have the lowest average similarity degrees with the remaining eastern Ede varieties (WLC1: 78%, WLC2: 66%). In contrast, all the other eastern Ede varieties have higher average similarity degrees with the remaining varieties of this cluster, on average 83% for WLC1 and 75% for WLC2. These findings substantiate the two subgroupings within the Eastern Ede cluster.

<99>

The distinct grouping of the Cabe varieties might be explained by the unique location of the Cabe communities. Like all eastern Ede varieties, the Cabe communities are located east of the Ouémé river and thus separated from the western Ede varieties and united with the eastern Ede varieties by this natural barrier. Geographically, however, the Cabe area is located at quite some distance from the remaining eastern Ede varieties; the southern Nago, Ije, and Yoruba (from Porto-Novo) communities of southeastern Benin are the closest to the Cabe area while the Nigerian Ede communities are still further away. At the same time though, the Cabe communities are geographically connected to the Western Ede varieties, namely through the International Main Road 2 (RNIE2) that connects Cotonou in southern Benin with Niamey (Niger). The RNIE2, coming from Parakou and going southwards, passes through the Cabe area and continues southwards to Dassa-Zoumé, the center of the Idaca language area. From Dassa-Zoumé, another International Main Road, the RNIE3/5, branches off eastwards. The RNIE3 continues eastwards, towards Togo, into the Ifè language area while, at Savalou, the RNIE5 branches off northwards to pass through the Ica language area and still further north through the northern Nago and Kura areas. Thus, the Cabe communities are located between the remaining eastern Ede varieties and the western Ede varieties while at the same time also being situated at quite some distance from either. This unique location may explain the overall lower degree of “Ede-ness” of the Cabe varieties as well as their overall grouping within the Eastern Ede cluster, in spite of the closer relationship to some of the western Ede varieties.

<100>

As with the western Ede varieties, both, the somewhat scattered placement of the eastern Ede varieties indicated by the MDS plots and the degree of variation found in the lexical similarity matrices, confirm the chaining pattern of the Ede varieties. The eastern Ede groups are situated geographically in a contiguous arrangement east of the Ouémé river, from the Cabe areas southwards to the Ije and southern Nago varieties and then eastwards to the southeastern Ede varieties of Nigeria.

Northern Ede

<101>

The Northern Ede cluster refers to the Mokole variety which is located in the far north of Benin, approximately 280 km north of the most northern communities of the Western Ede cluster of Kura and Cabe.

<102>

Both MDS plots indicate a distinct placement of the Mokole variety not shared by any of the other Ede varieties. Along the perceptual “Edeness” axis, Mokole is situated on the “-Ede” side of this axis, while it is placed on the Western side along the perceptual Eastern-Western Ede axis. This distinct placement indicates very low degrees of lexical similarity between Mokole and the remaining Ede varieties, a suggestion that appears to be supported by the respective average degrees of inter-group similarity for both word list computations, as displayed in Table 8.

Table 8: Average degrees of inter-group similarity between Northern Ede and the Western and Eastern Ede clusters

	Average inter-group similarity		Inter-group similarity ranges	
	WLC1	WLC2	WLC1	WLC2
North – West	54	44	51-56	39-48
North – East	52	40	47-57	33-46

<103>

Northern Ede, that is Mokole, shares very low degrees of lexical similarity with the remaining Ede varieties, both with the Western Ede (WLC1: 54%, WLC2: 44%) and the Eastern Ede cluster (WLC1: 52%, WLC2: 40%). The respective ranges of lexical similarity are rather small and confirm the overall low degree of similarity for Mokole with Western Ede (WLC1: 51-56%, WLC2: 39-48%) as well as with Eastern Ede (WLC1: 47-57%, WLC2: 33-46%).

<104>

These findings raise the question, whether and to what extent Mokole is in fact related to the remaining Ede and even the remaining Edekiri varieties. However, as already mentioned in Section 2.1, Capo (1989:280), Benin's National Language Commission (Commission Nationale de Linguistique 1983) and CENALA (1990) do include Mokole in their listings of Ede varieties. In contrast though, Gordon (2005) does not classify Mokole as an Ede variety but lists Mokole directly under the Edekiri node.

Discussion and conclusions

Clustering of the Ede varieties

<105>

One of the main objectives of the current study was to assess the clustering of the Ede speech varieties and to explore whether different sets of similarity judgment criteria lead to compatible conclusions as to the relative relationships and clustering of the investigated varieties.

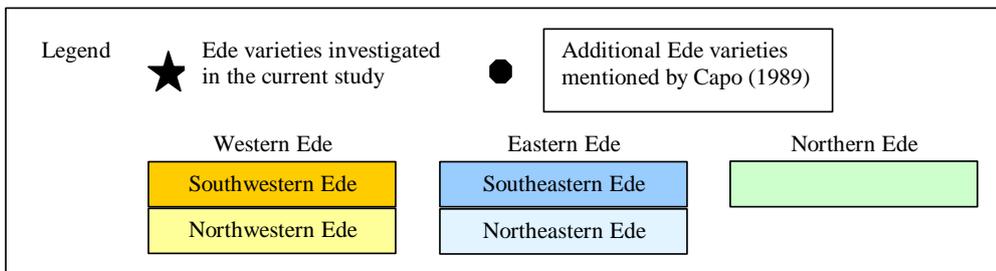
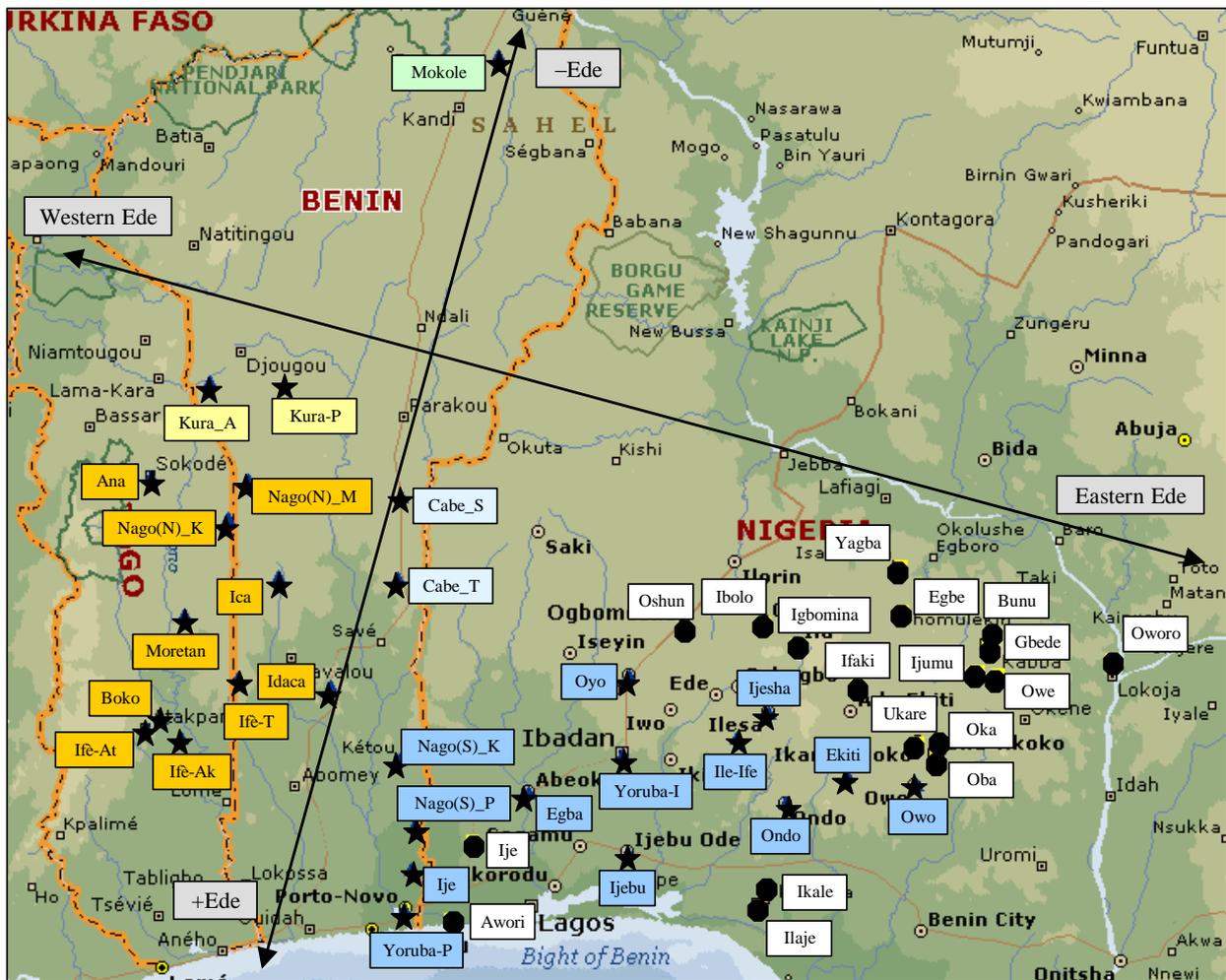
<106>

The different treatment of the elicited lexical Ede items according to two different sets of similarity judgment criteria resulted in different similarity judgment decisions for 110 glosses (67%) of the 163-item word list. These different decisions yielded significantly different lexical similarity percentages with consistently lower lexical similarity degrees for WLC1 which is based on a more rigorous set of similarity judgment criteria. At the same time though, the results of the correlation analysis indicated a significant linear relationship between the two computed word list similarity matrices, thus supporting the conclusion that use of either similarity judgment criteria sets would lead to similar clustering results.

<107>

The findings of the MDS analysis of both word list computations corroborate this conclusion, indicating three major Ede clusters, comprising the same subgroupings: western Ede referring to the southwestern and the northwestern Ede varieties, eastern Ede comprising the southeastern and northeastern Ede varieties, and northern Ede referring to the Mokole variety. (See Figure 6 for a map of the Ede language area, displaying the geographical locations of the Ede clusters and their components as indicated by the MDS findings; the superimposed arrows represent the two suggested MDS dimensions, that is the "Edeness" axis and the Western-Eastern Ede axis.)

Figure 6: Map of the Ede language area according to the MDS findings (based on Microsoft Corporation 2002)



<108>

In Table 9, the 28 investigated Ede varieties are classified into the identified Ede clusters and sub-groupings as indicated by the MDS findings for both word list computations, suggesting 12 western, 15 eastern, and 1 northern Ede varieties.

Table 9: Grouping of the Ede varieties according to the findings of MDS

Western Ede		Northern Ede	Eastern Ede	
Southwest	Ana (Sokodé)	Mokole	Egba	Southeast
	Boko		Ekiti	
	Ica		Ije	
	Idaca		Ijebu	
	Ifè (Akparè)		Ijesha	

Western Ede		Northern Ede	Eastern Ede	
	Ifè (Atakpamé)		Ile-Ife	
	Ifè (Tchetti)		Nago-South (Kétou)	
	Morétan		Nago-South (Pobè)	
	Nago-North (Kambolé)		Ondo	
	Nago-North (Manigri)		Owo	
Northwest	Kura (Awotébi)		Oyo	
	Kura (Partago)		Yoruba (Ibadan)	
			Yoruba (Porto-Novu)	
			Cabe (Savè)	Northeast
			Cabe (Tchaourou)	

<109>

For both word list computations, the suggested groupings concur for all 28 investigated Ede varieties, despite the significantly different lexical similarity percentages between both computations which are caused by the differing degrees of rigorousness of the underlying similarity judgment criteria.

<110>

For two of the Ede varieties, the MDS plots as well as the descriptive analysis of the computed similarity matrices indicate a possible alternative grouping: Kura from Awotébi and Kura from Partago. The distinct placement of the Kura varieties towards the “-Ede” side of the “Edeness” axis, together with the rather low lexical similarity degrees between the Kura varieties and southwestern Ede varieties, raise the question whether and to what extent these varieties are in fact part of the larger Ede cluster. An alternative grouping of the Kura varieties, together with Mokole within the Northern Ede cluster, does not, however, appear to be the most fitting option, given the low average degrees of lexical similarity between Kura and Mokole. On the other hand, if one defines Northern Ede rather loosely as a group of varieties that are somewhat related to the remaining Ede varieties although not closely, the grouping of the Kura varieties within the Northern cluster might after all constitute a promising alternative.

<111>

These findings substantiate the initial hypothesis that use of either set of similarity judgment criteria would indicate similar relative relationships and thus lead to similar clustering results regardless of the underlying different lexical similarity percentages.

<112>

The proposed clustering of Ede language varieties is only to a certain degree comparable to Capo’s (1989) previously-mentioned classification (see Section 2.1). This is mainly due to two facts: Capo’s (1989) classification considers a number of Nigerian Ede varieties that were not included in the current study, and (2) Capo’s (1989:281) classification is not exhaustive, i.e. only a selection of the Ede varieties listed by Capo (1989:279ff) are mentioned in this classification, leaving the classification of the remaining varieties somewhat uncertain. A comparison of both classifications and their components suggests that Capo’s (1989) classification provides a more detailed picture for the Nigerian Ede varieties, while the current study provides a more detailed classification for the Ede varieties of Togo and Benin (see Table 10).

Table 10: Major Ede groupings according to Capo (1989) and the current study

Capo (1989)	Ede varieties	Current study	
--	Kura	Western:	Northwestern
Southwest	Ana		Southwestern
	Cabe	Eastern:	Northeastern
	Ketu		Southeastern
Northwest	Egba, Oshun, Oyo		“ “
Central	Ife, Ijesha, Ekiti		“ “
Northeast	Gbedde, Ijumu, Yagba		--
--	Mokole		Northern

<113>

The Ede varieties of Togo and Benin are classified by Capo (1989) as southwest Ede (e.g. Ana, Cabe, and Ketu), whereas the current study classifies these varieties as northwestern (i.e. Kura), southwestern (e.g. Ana), northeastern (i.e. Cabe), or southeastern (e.g. Kétou-Nago) Ede varieties. With regard to the Nigerian Ede varieties, Capo (1989) distinguishes between northwest (e.g. Egba, Oshun, and Oyo), central (e.g. Ekiti, Ife, and Ijesha), and northeast (e.g. Gbedde, Ijumu, and Yagba) Ede. Meanwhile, the current study classifies these varieties as southeastern Ede varieties, since neither the MDS plots nor the computed lexical similarity degrees suggest a further subgrouping of the indicated southeastern Ede varieties. In addition, the findings of the current study suggest a third major grouping, Northern Ede, which comprises the Mokole variety.

<114>

Neither classification, though, is based on a thorough diachronic analysis. In the case of Capo's (1989: 283) classification this is mainly due to the fact that *"not enough information is available to arrive at reliable subclassifications"* on the basis of *"rigorous sound laws and specific innovations"*, while for the current study a synchronic approach was chosen, given the overall goal of SIL's larger Ede study of exploring the extensibility of already existing materials among the remaining Ede varieties. Therefore, both classifications need to be regarded with some reservation.

Recommendations for further sociolinguistic research

<115>

A second objective of this study was to explore whether different sets of similarity judgment criteria lead to compatible recommendations for more in-depth sociolinguistic research.

<116>

In applying SIL International's 70%-threshold, it is assumed that if lexical similarity is lower than 70%, the speech communities in question would not understand the other speech forms well enough to be able to benefit from the same language development efforts. Therefore, more in-depth research (i.e. dialect intelligibility testing) assessing the extensibility of already existing materials would be of a lower priority. In contrast, if lexical similarity is higher than 70%, more in-depth research is recommended to investigate whether the respective speech communities can understand the other speech forms well enough to be able to benefit from the same literacy efforts. In light of already existing literature in Ifè and in Yoruba, further research among the Ede communities would therefore need to give the highest priority to those Ede speech communities that belong to the same language group (in terms of the 70%-threshold) but are not components of a cluster or subgrouping with ongoing language development efforts.

<117>

The findings for the current study indicate that the application of different similarity judgment criteria results in different conclusions as to whether the investigated Ede varieties belong to the same or distinct language groups. As far as the second level of the Ede clustering hierarchy is concerned, though,

these differing conclusions do not result in different recommendations concerning more in-depth research. The WLC2 findings suggest that the three identified Ede clusters belong to distinct language groups. Dialect intelligibility testing to assess the extensibility of already existing materials to any of the other clusters would, therefore, not be of high priority. In contrast, the WLC1 findings indicate that the Western and Eastern Ede clusters belong to the same larger language group. However, since both constitute clusters with ongoing language development efforts, more in-depth research assessing whether existing literature in one cluster would appropriately extend to use by the communities of the other cluster would, therefore, not be of high priority.

<118>

As regards more in-depth research (i.e. intelligibility testing) within each of the identified three major Ede clusters, the findings for both word list computations lend themselves to the following recommendations.

Western Ede

<119>

In light of existing literature in Ifè, a southwestern Ede variety, further research among the western Ede communities would need to give the highest priority to those western Ede speech communities that are not components of the southwestern subgrouping, that is the northwestern Ede or Kura varieties.

<120>

Such research would need to assess whether the Kura communities could benefit from Ifè development efforts, or whether an additional development program for the Kura varieties would be beneficial. The findings indicate, however, that in terms of the 70%-threshold the Kura varieties do not belong to the same language group as the southwestern Ede varieties. In consequence, more in-depth research with regard to the extensibility of Ifè materials among the Kura communities would not be assigned high priority. In contrast, the MDS findings appear to suggest that the Kura varieties group within a larger Western Ede cluster, together with the southwestern Ede varieties. These findings would warrant more in-depth research among the Kura varieties to assess the linguistic and socio-linguistic status of Kura and to explore the potential for Ifè and Yoruba literature extensibility.

<121>

More in-depth research among the southwestern Ede communities would need to investigate whether existing Ifè literature would appropriately extend to use by all the other southwestern Ede communities. With regard to Manigri-Nago though the WLC2 findings indicate that, in terms of the 70%-threshold, Manigri-Nago does not belong to the same language group as the remaining southwestern Ede varieties. This in turn leads to the conclusion that more in-depth research among the Manigri-Nago speech variety would be of lesser priority. Further research among the southwestern Ede would also need to verify the *Ethnologue*'s (Gordon 2005) classification of Ifè as an Edekiri but not as an Ede language.

<122>

Furthermore, given that Yoruba is one of Benin's official languages for large-scale non-formal adult education, research among the western Ede varieties of Benin would preferably also explore whether existing Yoruba materials would appropriately extend to use by the Western Ede communities, again giving special attention to the Kura and Manigri-Nago communities.

Eastern Ede

<123>

Given the ongoing Yoruba language development efforts, the highest priority for further research would need to be given to the Cabe communities which among all eastern Ede varieties have the lowest average similarity degrees with the remaining eastern Ede varieties. Such research would need to assess whether and to what extent these communities could benefit from the existing literacy efforts in Yoruba and/or – given their unique geographical location between the eastern and western Ede communities – in Ifè or whether additional language-based development for the Cabe varieties would

be beneficial. The WLC2 findings indicate, however, that, in terms of the 70%-threshold, the Cabe varieties do not belong to the same language group as either the southeastern or the western Ede varieties. Thus, more in-depth research among the Cabe communities would be of a lesser priority.

<124>

More in-depth research among the eastern Ede varieties would also need to investigate whether the already existing Beninese Yoruba literature could appropriately be used by the remaining eastern Ede varieties of Benin, namely Ije and southern Nago (from Kétou and Pobè). Given the rather distant geographical location of these communities from the Ifè area and further given that Yoruba and not Ifè is one of Benin's official languages for non-formal education, further research regarding the extensibility of Ifè literature among the southeastern Ede communities would be of lesser priority. In addition, research among the eastern Ede varieties of Nigeria would need to assess whether the already existing Nigerian Yoruba literature would appropriately extend to use by all Nigerian Ede communities.

Northern Ede

<125>

The findings of the current analysis suggest that, in terms of the 70%-threshold, Mokole does not belong to the same language group as the western or eastern Ede varieties and, therefore, more in-depth research regarding Ifè and Yoruba literature extensibility would not be of a high priority. However, given Mokole's classifications as an Ede variety by sources such as Capo (1989:280), Benin's National Language Commission (Commission Nationale de Linguistique 1983) and CENALA (1990), additional research would still appear to be beneficial. Such research would need to assess whether and to what extent Mokole should in fact be classified as an Ede variety and what its linguistic relationship is to the remaining Edekiri varieties. Part of this research would also need to investigate whether and to what extent the Mokole communities would benefit from ongoing language development efforts in Yoruba and/or Ifè, both in terms of accessibility (i.e. comprehension) and acceptability (i.e. affinity), or whether an additional development program for the Mokole varieties would be beneficial.

Conclusions

<126>

For the Ede data set, the findings of the current study suggest that the consistent application of different similarity judgment criteria sets yields compatible conclusions as to the relative relationships of the investigated varieties which in turn leads to similar clustering results, despite the underlying different lexical similarity percentages. However, these clustering results do not necessarily lend themselves to the same recommendations for more in-depth sociolinguistic research, when interpreted in terms of SIL International's 70%-threshold. These differing recommendations, in turn, raise the question as to which of the underlying similarity judgment criteria sets is more appropriate for the Ede language situation.

<127>

Given these ambiguities, the approach chosen for the analysis of the Ede data set was to employ the techniques previously suggested for the analysis of the Gbe data set (Kluge 2005) which provided an informing case study regarding the effects of different lexical similarity judgment criteria sets. Thus, a combination of techniques, namely correlation analysis and multidimensional scaling, was applied which explored the relative relationships and clustering of the Ede varieties, rather than fixating on the suggested threshold level of 70% per se.

<128>

Given the limited scope of the larger Ede study of assessing literacy extensibility, this combination of different techniques was found to provide a convenient approach to the analysis of the rather large Ede data set. In particular, this methodology facilitated the synchronic analysis of the elicited Ede data set without requiring a thorough morphemic analysis which would have been, although preferable, more time consuming than was feasible for this study. Moreover, this approach facilitated the synchronic analysis of lexical items with polymorphemic word-structure for which an informed, commonly agreed-upon approach is still lacking. Finally, the comparison of the computed MDS plots, derived

from two similarity judgment criteria sets which differ in rigorousness, allowed to compare and verify the suggested groupings and their components.

<129>

In summary, the current study substantiates the conclusions of the Gbe analysis (Kluge 2005), in that the results for both word list computations indicate compatible conclusions as to the relative relationships of the investigated varieties which in turn lead to similar clustering results. In contrast to the Gbe study, though, the current analysis shows that these clustering results do not necessarily lend themselves to the same recommendations regarding more in-depth sociolinguistic research, when interpreted in terms of an absolute threshold. The indicated ambiguities thus substantiate the usefulness of focusing on the relative, rather than absolute, lexical relationships and clustering of the investigated varieties by means of multidimensional scaling in establishing recommendations for more in-depth sociolinguistic research.

References

Aguidi, R. 1992

Signification du nom individuel chez les Nagots de Bantè. Mémoire de maîtrise en linguistique. Abomey-Calavi: Université Nationale du Bénin. ms.

Akinkugbe, Olufemi Odutayo 1978

A comparative phonology of Yoruba dialects, Isekiri, and Igala. Ph.D. Dissertation. University of Ibadan, Nigeria

Baloubi, Désiré 2004

'The Idaacha communities in the Republic of Benin: Juggling for linguistic and ethnic identity.' In: Ashley, Leonard R.N. and Wayne H. Finke (eds.) *Language and identity: Selected papers of the International Conference, October 2-5, 2002 sponsored by the American Society of Geolinguistics Baruch College (CUNY)*. East Rockaway, NY: Cummings & Hathaway, pp.49-62

Blair, Frank 1990

Survey on a shoestring: A manual for small-scale language surveys. Dallas: Summer Institute of Linguistics and University of Texas

Boëthius, Hélène 1983

Description préliminaire des phonèmes segmentaux et types de syllabes ifè. In: Nicole, Jacques (ed.) *Études linguistiques préliminaires dans quelques langues du Togo*. Lomé: Société Internationale de Linguistique, pp.110-37

Boëthius, Hélène 1987

'Basic mood in Ifè.' *Journal of West African Languages* 17,2:43-70

Boëthius, Hélène 1991

Writing tone in Ifè. Lomé: Société Internationale de Linguistique. ms.

Brye, Edward and Elizabeth Brye 2004

'Intelligibility testing survey of Bebe and Kemezung and synthesis of sociolinguistic research of the Eastern Beboid cluster.' In: *SIL Electronic Survey Reports*, 2004-011. Dallas: SIL, <http://www.sil.org/silesr/2004/silesr2004-011.pdf> (29.03.2008)

Capo, Hounkpati B.C. 1989

'Defoid.' In: Bendor-Samuel, John (ed.) *The Niger-Congo languages*. Lanham: The University Press of America, pp.275-290

Centre National de Linguistique Appliquée (CENALA) 1990

République du Bénin – Carte linguistique. (Carte établie 1989). Cotonou: CENALA

Commission Nationale de Linguistique du Bénin (CNL du Bénin) 1983

Atlas sociolinguistique du Bénin. Abidjan: Agence de Coopération Culturelle et Technique, and Institut de Linguistique Appliquée

- Dettweiler, Stephen and Sonia Dettweiler 2003
 ‘Sociolinguistic survey of the Duka (Hun-Saare) people.’ In: *SIL Electronic Survey Reports*, 2004-004. Dallas: SIL International, <http://www.sil.org/silesr/2003/silesr2003-014.pdf> (29.03.2008)
- Devine, Patricia 1996
 Personal communication (email). September 5, Atakpamé, Togo
- Devine, Patricia 2004
 Personal communication (email). November 17, Pittsburgh, USA
- Dieu, Michel and Patrick Renaud (eds.) 1983
Atlas linguistique du Cameroun. ALCAM. Yaoundé, Cameroon: Délégation Générale à la Recherche Scientifique et Technique
- Direction de l’Alphabétisation, République du Bénin, Ministère de la Culture et des Communications (DA) 1992
Séminaire nationale sur la redéfinition des objectifs et stratégies d’alphabétisation et d’éducation des adultes. Rapport final. Cotonou: Presse de Onepi
- Efron, Bradley and Robert J. Tibshirani 1993
An introduction to the bootstrap. New York: Chapman and Hall
- Garrett, Peter, Nikolas Coupland, and Angie Williams 1999
 ‘Evaluating dialect in discourse: Teachers’ and teenagers’ responses to young English speakers in Wales.’ In: *Language in Society* 28:321-54
- Grimes, Joseph E. 1988
 ‘Correlations between vocabulary similarity and intelligibility.’ *Notes on Linguistics* 41:19-33
- Gudschinsky, Sarah C. 1955
 ‘Lexico-statistical skewing from dialect borrowing.’ *International Journal of American Linguistics* 21:138-149
- Gudschinsky, Sarah C. 1956
 ‘The ABC’s of lexicostatistics (glottochronology).’ *Word* 12:175-210
- Hair, Joseph F., Rolph E. Anderson, Ronald L. Tatham, and William C. Black 1998
Multivariate data analysis. London: Prentice-Hall International, Inc.
- Harrison, Byron, Annette Harrison, and Michael J. Rueck 1999
 ‘Southern Songhay speech varieties in Niger: A sociolinguistic survey of the Zarma, Songhay, Kurtey, Wogo, and Dendi peoples of Niger.’ In: *SIL Electronic Survey Reports*, 1999-004. Dallas: SIL International, <http://www.sil.org/silesr/1999/004/zarmarpt4.pdf> (29.03.2008)
- Hochstetler, J. Lee, Jude A. Durieux, Evelin I. K. Durieux-Boon 2004
 ‘Sociolinguistic survey of the Dogon language area.’ In: *SIL Electronic Survey Reports*, 2004-004. Dallas: SIL International, <http://www.sil.org/silesr/2004/silesr2004-004.pdf> (29.03.2008)
- International Language Assessment Conference 1990
 ‘Language assessment criteria: Conference recommendations.’ In: Kindell, Gloria E. (ed.) *Proceedings of the Summer Institute of Linguistics International Language Assessment Conference, Horsleys Green, 23-31 May 1989*. Dallas: Summer Institute of Linguistics, pp.27-29
- Klaver, Marquita 1987
 The Ifè kinship system. Lomé: Société Internationale de Linguistique. ms.
- Klaver, Marquita 1995
 ‘Two Ifè particles and their discourse function.’ *Journal of West African Languages* 25,2:107-13
- Klaver, Marquita 1999
 ‘A look at Ifè connectives.’ In: Loos, Eugene E. (ed.) *Logical relations in discourse*. Dallas: Summer Institute of Linguistics, pp.67-72

- Kluge, Angela 1999
Sociolinguistic survey of the Ede language communities of Benin and Togo: General overview. Cotonou: SIL International. ms.
- Kluge, Angela 2005
'A synchronic lexical study of Gbe language varieties: The effects of different similarity judgment criteria.' *Linguistic Discovery* 3,1:22-53,
<http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/article/281> (29.03.2008)
- Kohler, Anna 1983
'Le ton en ifè.' In: Nicole, Jacques (ed.) *Études linguistiques préliminaires dans quelques langues du Togo*. Lomé: Société Internationale de Linguistique, pp.138-156
- Kruskal, Joseph B. and Myron Wish 1978
Multidimensional scaling. Beverley Hills: Sage Publications
- Manly, Bryan F.J. 1986
Multivariate statistical methods: A primer. London, New-York: Chapman and Hall
- Meyers, Jim, Heather Meyers, Ben Rice, and Susan Rice 2003
'Report on the Seruyan river surveys in central Kalimantan: Kohin, Keninjal and Sebaun (Dohoi) languages.' In: *SIL Electronic Survey Reports*, 2003-017. Dallas: SIL International,
<http://www.sil.org/silesr/2003/silesr2003-017.pdf> (29.03.2008)
- Microsoft Corporation 2002
Microsoft Encarta reference library. Microsoft Corporation
- Ministère du Plan et de la Restructuration Economique du Bénin. Institut National de la Statistique et de l'Analyse Economique. Bureau Central du Recensement 1994
Deuxième recensement général de la population et de l'habitation (Février 1992). Volume II: Analyses des résultats. Tome 3: Caractéristiques socio-culturelles et économiques. Cotonou: Ministère du Plan et de la Restructuration Economique du Bénin
- Norušis, Marija J. 1993
SPSS for Windows: Professional statistics. Release 6.0. Chicago: SPSS Inc.
- Odoun Kouyomou, Dieudonné 1986
Phonologie de la langue idaasha. Mémoire de maîtrise en linguistique. Abomey-Calavi: Université Nationale du Bénin. ms.
- Odoun Kouyomou, Dieudonné 1989
Apprenons à lire et à écrire en langue idàáshà. Cotonou: CENALA
- Odoun Kouyomou, Dieudonné 1990
Petit lexique bilingue français-idàáshà sur la poste, la banque, la santé, l'agriculture (terminologie). Cotonou: CENALA
- Odoun Kouyomou, Dieudonné 1991a
Apprenons à transcrire la langue ifè. Cotonou: CENALA
- Odoun Kouyomou, Dieudonné 1991b
Èdè idàáshà kiko: Comment transcrire l'idàácà. Cotonou: CENALA
- Odoun Kouyomou, Dieudonné 1991c
L'énoncé et les prédicatifs nominaux en idàáshà. Cotonou: CENALA
- Odoun Kouyomou, Dieudonné 1991d
Petit lexique français-idàáshà sur les animaux et oiseaux domestiques, les oiseaux sauvages et la faune du Bénin. Cotonou: CENALA
- Odoun Kouyomou, Dieudonné 1991e
Terminologie français-idàáshà sur les mesures de capacité, de poids, de distance, de temps, de volume. La numération et l'énumération. Cotonou: CENALA

- Parrinder, Edward Geoffrey 1947
 'Yoruba-speaking people in Dahomey.' *Africa* 17,2:122-129
- Probst, Ulrich 1992
 Multimorphemic words in lexicostatistics: An evaluation on use and limitations of different approaches. Paper presented at the SIL Survey Course, Horsleys Green, UK. ms.
- R-Project, no date
The R project for statistical computing, <http://www.r-project.org/> (01.02.2008)
- Rueck, Mike and Tim Jore 2003
 'The sociolinguistic situation of the Sam people.' In: *SIL Electronic Survey Reports*, 2003-019. Dallas: SIL International, <http://pnglanguages.org/silesr/2003/silesr2003-019.pdf> (01.02.2008)
- Sanders, Arden G. 1977
 'Guidelines for conducting a lexicostatistics survey in Papua New Guinea.' In: Loving, Richard E. and Gary F. Simons (eds.) *Language planning and survey techniques. Workpapers in Papua New Guinea Languages* 21. Ukarumpa: Summer Institute of Linguistics, pp.21-41
- Saussure, Ferdinand de 1959
Course in general linguistics. London: Peter Owen Limited
- Savage, Dale 1999
 'Understanding correlation.' In: Casad, Eugene H. (ed.) *Windows on bilingualism*. Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington, pp.117-146
- Scholfield, Phil 1991
 'Statistics in linguistics.' *Annual Review of Anthropology* 20:377-393
- Schooling, Stephen J. 1981
A linguistic and sociolinguistic survey of French Polynesia. Hamilton, New Zealand: Summer Institute of Linguistics
- SIL International 2007
Complete index of SIL publications: Togo. Dallas: SIL International, http://www.ethnologue.com/show_country_bibl.asp?name=TG (01.02.2008)
- Simons, Gary F. 1979
Language variation and limits to communication. Technical Report 3. Ithaca: Department of Modern Languages and Linguistics, Cornell University (reprinted 1983 by Summer Institute of Linguistics, Dallas)
- Simons, Gary F. 1977
 'The role of purpose and perspective in planning a language survey.' In: Loving, Richard E. and Gary F. Simons (eds.) *Language planning and survey techniques*. Ukarumpa, PNG: Summer Institute of Linguistics, pp.9-20
- SPSS Inc. 2006
SPSS for Windows: Release 15.0. SPSS, Inc., 1989-2006
- Swadesh, Morris 1955
 'Toward greater accuracy in lexicostatistical dating.' *International Journal of American Linguistics* 21:121-137
- Williamson, Kay 1989
 'Niger-Congo overview.' In: Bendor-Samuel, John (ed.) *The Niger-Congo languages*. Lanham: The University Press of America, pp.3-45
- Wimbish, John 1989
WordSurv: A program for analyzing language survey word lists. Dallas: Summer Institute of Linguistics
- Woods, Anthony, Paul Fletcher, and Arthur Hughes 1986
Statistics in language studies. Cambridge: Cambridge University Press

-
- ¹ I would like to thank Ted G. Bergman, Robert C. Cochran and Ramzi Nahhas of SIL International for their helpful comments on this paper. Nahhas, in particular, discussed various aspects of paired t-test and correlation analysis with me. I would also like to thank the two anonymous referees for their most helpful critiques.
- ² In general, all language names are spelled according to the Ethnologue (Gordon 2005).
- ³ In addition, Gordon (2005) lists Lucumi, “a secret language used for ritual by the Santeria religion” in Cuba.
- ⁴ For ease of reference, from this point onward the term Ede also subsumes the Yoruba and Ifè varieties even though the Ethnologue (Gordon 2005) classifies them directly under the Edekiri node and not as Ede varieties.
- ⁵ Nonvocalic segments are either exact matches or else they differ by only one phonological feature and this difference is attested in three pairs, and vowels differ by only one phonological feature.
- ⁶ Nonvocalic segments differ by only one phonological feature but are not attested in three pairs; vowels differ by two or more phonological features.
- ⁷ Thus, if A and B share 50% lexical similarity and the range of error is +/-25% then the actual lexical similarity could be anywhere from 25-75%. And if A and C share 68% lexical similarity and the range of error is +/-7% then the actual lexical similarity would be anywhere from 61-75%. In both cases ‘75%’ refers to ‘upper range of error’ and, based on SIL’s criterion, intelligibility testing between A and B and A and C would need to be considered.