**Workshop 'Language Analysis in the determination of national origin'**

21st of June, Amsterdam

**Programme**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00-9:30</td>
<td>Coffee and Opening</td>
</tr>
<tr>
<td>9:30-10:00</td>
<td>T. McNamara (University of Melbourne)</td>
</tr>
<tr>
<td></td>
<td><em>Language analysis: A question of validity</em></td>
</tr>
<tr>
<td>10:00-10:30</td>
<td>V.A. de Rooij (University of Amsterdam)</td>
</tr>
<tr>
<td></td>
<td><em>(Im)Possibilities of language analysis as a means of establishing the nationality of asylum seekers</em></td>
</tr>
<tr>
<td>11:00-11:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>11:30-12:00</td>
<td>T. Cambier-Langeveld &amp; A. Samson (Office for Country Information and Language Analysis GCKAO, Immigration and Naturalization Service (the Netherlands))</td>
</tr>
<tr>
<td></td>
<td><em>Language analysis: how to include both linguistic expertise and native competence, and why</em></td>
</tr>
<tr>
<td>12:00-12:30</td>
<td>I. Sawicka (Institute of the Slavic Philology)</td>
</tr>
<tr>
<td></td>
<td><em>Recognition national origin by phonetic information</em></td>
</tr>
<tr>
<td>12:30-12:50</td>
<td>Discussion</td>
</tr>
<tr>
<td>12:50-13:50</td>
<td>Lunch</td>
</tr>
<tr>
<td>13:50-14:20</td>
<td>E.M. Bergman (Georgetown University)</td>
</tr>
<tr>
<td></td>
<td><em>Spoken Arabic and the Language Analyst</em></td>
</tr>
<tr>
<td>14:20-14:50</td>
<td>R. Thelwall (University of Calgary)</td>
</tr>
<tr>
<td></td>
<td><em>The Argument from Silence? Diglossia and the identification of Sudan Arabic dialects</em></td>
</tr>
<tr>
<td>14:50-15:20</td>
<td>Shishir Bhattacharja (University of Montreal)</td>
</tr>
<tr>
<td></td>
<td><em>Outlines of Forensic linguistic analysis for determining national origin</em></td>
</tr>
<tr>
<td>15:20-15:40</td>
<td>Coffee break</td>
</tr>
<tr>
<td>15:40-16:10</td>
<td>J. Ndayiragije (University of Toronto)</td>
</tr>
<tr>
<td></td>
<td><em>Language Analysis and Microvariation: A Case Study in Bantu Relatives</em></td>
</tr>
<tr>
<td>16:10-16:40</td>
<td>M.A. Bah (University of Charleston)</td>
</tr>
<tr>
<td></td>
<td><em>The problems of language determination and citizenship among ethnic groups spread across national political boundaries: the case of Pular and Fulbe in the Mano river tri- states of Sierra Leone, Liberia and Guinea</em></td>
</tr>
<tr>
<td>16:40-17:00</td>
<td>Discussion</td>
</tr>
<tr>
<td>17:00</td>
<td>Closure and Drinks</td>
</tr>
</tbody>
</table>

**Abstracts**

*The problems of language determination and citizenship among ethnic groups spread across national political boundaries: the case of Pular and Fulbe in the Mano river tri- states of Sierra Leone, Liberia and Guinea*

**M.A. Bah** (University of Charleston)

Colonial boundaries established by European colonial powers in Africa remain the worst legacy left behind by the colonizers. However, African nationalism and the current concept of the nation state with fifty-three African countries only exist because of the established European colonial boundaries. One of the major current sources of conflict within the continent is the delimitation of the boundaries, dividing families into two or three countries. In addition to the division of families, larger ethnic groups are also divided into multi nation states. Ultimately, there are fixed boundaries dividing territories, but not people. This phenomenon negates and challenges the concept of patriotism. Groups are often more passionate about their ethnic rather than their national
commitments and origin. Added to this complexity is the use of ethnic languages and especially formal state languages and lingua francas. In the case of the Mano River states of Sierra Leone, Liberia, and Guinea many languages are spoken in all three countries by the same ethnic groups. There are considerable number of Kissis, Mandingos, and Fulbe in all three countries. There are also Mendes, Vais, Golas, Krus, Bassas in both Sierra Leone and Liberia. The Susus, Korankos, and Yallunkas are found in both Sierra Leone and Guinea. The Loma, Mano, and Gio are are in both Liberia and Guinea. Thus, all the languages of the above groups are spoken in all three countries. Any objective and sincere determination of citizenship in these cases are often too difficult.

The Fulbe and their language present a typical example of these difficulties. The Fulbe are found all over West Africa and in some parts of central and east Africa. They often speak the same language Puular with some variations depending on their geographical locations. They are also usually Muslims with some serious attachments to their culture and claim to have been agents of the spread of Islam in West Africa. They are usually traders and Islamic teachers who generally resist inter marriages with local ethnic groups especially non-Muslims. Because of colonial experience, the Fulbe have picked up vocabularies from the languages of the colonizers. They have also over time picked up languages and considerable vocabularies from their neighboring ethnic groups. I plan to discuss the varieties of Puular and the difficulties involved in determining Fulbe origin depending on colonial or indigenous borrowed vocabularies. I would also like to discuss Fulbe's ability or inability to properly acquire the languages and cultures of their neighbors. Another major difficulty about Fulbe is the bias against them for being spread all over West Africa, their unshakeable commitment to Islam, and their success in business. The Fulbe in the Mano River states were often marginalized, but because of success in business and growth in population, the case is no more the same. Finally, an understanding of the Fulbe, their language and, philosophy will help all interested parties minimize misunderstandings and reduce the stereotypes about the Fulbe.

Spoken Arabic and the Language Analyst

E.M. Bergman (Georgetown University)

Spoken Arabic poses a number of challenges for the language analyst. First is the nature of Spoken Arabic. It is the language variety of primary discourse, of daily life and personal interaction. It is distinct – but never separate – from Modern Standard Arabic (or Literary Arabic), which is the variety of secondary discourse, used in institutional settings and the written word. Where Modern Standard Arabic is standardized, Spoken Arabic is not. Spoken Arabic differs in identifiable ways from country to country. The language analyst faces a minimum of 22 varieties of Arabic, one for each nation in the League of Arab States. Even within a single country, however, Spoken Arabic can differ from region to region, from city to city, even from one village to the next. As linguistic differences exist in phonology, morphology, syntax, and lexicon, the language analyst may face hundreds of possible linguistic variables and varieties.

Demographics further complicate the situation. The term “the Arab world” masks the variety of ethnicities within the territory it covers. A relatively small country like Syria, for example, shelters 15 different language varieties. A Syrian speaker of Arabic may, thus, have Armenian, Circassian, or Kurdish as a home variety. These varieties are likely to influence a speaker’s Arabic in ways that are not widely known outside the speaker’s community. And when the speaker’s community changes, the speaker’s language is more than likely to change with it. The Arab world has known a number of mass population displacements in recent years, in addition to the effects of urbanization. Population movements within and from Iraq, to name only the most current example, result in individuals living far from their community or country of origin. The effect on the Arabic they speak is as yet unknown.
With all of these complications, how can a language analyst produce reliable results for Arabic? This paper will suggest some strategies and make recommendations for the language analyst and for those who recruit and work with analysts.

Outlines of Forensic linguistic analysis for determining national origin

S. Bhattacharja (University of Montreal)

By ‘forensic linguistic analysis’ I mean ‘a particular process that is aimed either to determine or to verify the identity of an individual on the basis of the (i) phonological, (ii) morphological, (iii) syntactical, (iv) lexical and (v) stylistic features of an oral or a written text produced by him.’ ‘Nation’ generally means ‘a group of people with some common physical features, language and culture, not necessarily living in the same state.’ To start with, a forensic linguistic analyst must be well aware of, or should preferably have in hand a detailed phonological, morphological and syntactic description of the language of the text in question. Such a description may have been done from the point of view of any elegant theory available in each of these core domains of linguistics. On the basis of a comparison between the pertinent linguistic peculiarities of the text and the linguistic features of the language it is possible to speculate on the national origin of its speaker, or in case of written text, its writer. In this article, I will show how such an analysis can be done and also indicate some of its shortcomings. However, I will argue that the shortcomings of forensic linguistic analysis in determining national origin are, to a large extent, due to the fact that linguistic identity and national identity may not be inclusive of each other.

Language analysis: how to include both linguistic expertise and native competence, and why

T. Cambier-Langeveld & A. Samson (Office for Country Information and Language Analysis GCKAO, Immigration and Naturalization Service (the Netherlands))

In 2004 the Language and National Origin Group published guidelines for the practice of language analysis for the determination of origin in refugee cases. In these guidelines, it is argued that linguistic expertise should always be involved in such language analyses. Restrictions and limitations of language analysis are also listed and explained.

We agree with the main points of the guidelines and we are actively involved in making our customers aware of the limitations of language analysis as a means to determine the origin of asylum seekers. We also recognise the dangers of language analyses by native speakers without (the involvement of) linguistic expertise, and would strongly disapprove of such practice. However, there are also risks attached to language analyses by linguistic experts without (the involvement of) native competence of the language involved. These latter risks are not addressed in the guidelines, but will be addressed in this presentation.

Beside the issue of the qualifications of the person performing the analysis, there is another issue which has come to play an increasingly large role in the way we organise and control our use of language analysis. This concerns the testing of the analyst’s ability to distinguish between relevant language varieties and also between authentic and nonauthentic speech features.

In this presentation, we will discuss one workable solution in which all the requirements mentioned above are met. This is a construction where the expertise of a linguist and the language competence of a native speaker are both involved yet not united in one person. We will explain in the presentation how this is implemented in practice and what the advantages and disadvantages are of this construction.
(Im)Possibilities of language analysis as a means of establishing the nationality of asylum seekers

Vincent A. de Rooij (University of Amsterdam)

Language analysis reports commissioned by the Dutch Immigration Service (IND) often do not meet basic requirements of sound scientific research. It is worrying to notice that IND experts seem to lack essential knowledge of procedures of linguistic analysis. But even if they would have this knowledge, the problem remains that interviews between asylum seekers, immigration officers and interpreters, have characteristics that raise questions about the possibility of making reliable judgements of national origin on the basis of language analysis in general. This paper will offer a critical evaluation of some of these features and suggest an alternative method of data collection which does not suffer from the shortcomings of the procedure used until now. In the proposed procedure for data collection, the asylum seeker's spontaneous speech is recorded in a familiar setting. This is thought to result in more reliable data than data collected in formal settings in which asylum seekers may display accommodation behaviors that are hard to interpret. Even data, however, collected in settings and situations familiar to the asylum seeker, may not always be suited for establishing speaker's origin or nationality. In some cases language experts (IND experts or independent experts doing contra-expertises) should be willing to admit that scientifically reliable judgements of origin through language analysis is impossible.

Language analysis: A question of validity

T. McNamara (University of Melbourne)

Validity theory in language assessment provides a possible framework for considering the place of language analysis in the adjudication of the legitimacy of the claims of asylum seekers.

The paper provides an outline of validity theory in relation to language assessment, drawing on the work of Messick and Kane, and considers the following aspects as relevant to language analysis in the case of asylum seekers:

1. the adequacy of the construct, which in this case is sociolinguistic in character (involving varieties of language spoken in the area in question, contact with other varieties, varieties encountered on the journey to the place where the application was filed, etc);
2. the inevitable uncertainty of the inferences to be drawn from the evidence available in the language sample, and the probabilistic nature of conclusions;
3. the adequacy of the sample used as the basis for the conclusions reached;
4. the procedures used to evaluate the sample;
5. the social values implicit in the assessment;
6. the social consequences of the assessment.

It will be argued that this framework offers a sound and useful basis for considering the appropriate and permissible roles for the use of language evidence in asylum seeker procedures. Examples from the Australian context will be considered in relation to each of the points raised.
Language Analysis and Microvariation: A Case Study in Bantu Relatives

J. Ndayiragije (University of Toronto)

This talk shows how a close investigation of syntactic and phonological properties of relative clauses in Bantu may shed light on the national origin of asylum seekers. I examine relativization strategies in four languages: CiShona (Zimbabwe, Botswana, and Zambia), Bemba (Zambia, Democratic Republic of Congo -DRC), Kirundi (Burundi, DRC, Rwanda, Tanzania), and Kinyarwanda (Rwanda, Burundi, DRC, Tanzania).

Cheng (2006) identified two relativization strategies in Bantu languages: the Possessive strategy (1a-b), from ChiShona, and the Demonstrative strategy (2c), from Bemba:

(1) a. ndi mi za-va-nótaúra 10language 10Rel-2SM-speak
"Languages that they speak"

b. múunú wa-ku-húúva 1person 1REL-INFL-have.problems 'a person (who is) in trouble'

(2) a. Umulembo a-ka-belenga ibuku 1boy 1SM-FUT-read 5book
'The boy will read the book'

b. Umulembo u-u-ka-belenga ibuku SUBJ Relative 1boy 1REL-1SM-FUT-read book 'The boy who will read the book'

c. Ibuku i lyo umulembo a-ka-belenga OBJ Relative 5book 5Rel 1boy 1SM-FUT-read
'The book that the boy will read'

Cheng and Kula (2006) also found that phonology matters, witness tone lowering on the subject marker (SM) in Cishona (3b), and Bemba (4b):

(3) a. ba-kafúndisha á-bá-léé-lolesha pansé ni ba-Mutale. 2PFX-teacher 2REL-2SM-TNS-look 16outside COP 2PFX-Mutale
'The teacher who is looking outside is Mr Mutale'

b. ba-kafúndisha ba-léé-lolesha pansé ni ba-Mutale.
2PFX-teacher 2SM-TNS-look 16outside COP 2PFX-Mutale
'The teacher who is looking outside is Mr Mutale'

(4) a. murúmé á-ca-tongá nyika 1man 1SM-FUT-rule country
'The man will run the country'

b. murúmé á-ca-tongá nyika SUBJ Relative 1man 1SM-FUT-rule country 'The man who will run the country'

On theoretical grounds, (1-4) raise interesting questions about the relevance of tonology in syntactic computation, and the parametric source of the Possessive-Demonstrative asymmetry.
This talk answers to those questions by deriving the tonal change and the Demonstrative-Possessive asymmetry from a general PF (phonological) requirement of complementizer lexicalization in those languages, similar to V-second effect in Germanic languages.

Evidence supporting the proposal comes from other syntactic similarities and contrasts found in other Bantu languages: Kirundi and Kinyarwanda, in which Possessive and Demonstrative strategies are banned where they are required in CiShona and Bemba, but tonal change is mandatory. Methodological implications are drawn on the relevance of syntactic and phonological features of asylum seekers’ speech in the determination of their national origin.

**Recognition national origin by phonetic information**

*I. Sawicka* (Institute of the Slavic Philology)

The most frequent phonetic values of a language are not features belonging to particular segments but to whole classes of sounds in given contexts. They give a common impression of how a language sounds and cause us to describe the pronunciation in general terms of sensory perception such as, for example, hard, soft, tense, nasal, aspirated, monotonous, singsong, buzzing, clear, non-distinct, well-marked, etc. Such a general character is inherent to particular languages, it is relatively stable, it is not subject to interference, it is difficult to learn and difficult to get rid of. We observe it in pronunciation of foreigners and in the way in which loanwords are adapted. It is connected to phonemic values as well as to type of combinatory allophony and the way in which segments combine on the syntagmatic axis. These values should be known and described but often they are not, especially when a feature is regular but not obligatory. Very often linguists and even phoneticians which are native speakers are not even conscious of them. Descriptions concentrate first of all on phonemically relevant features and on obligatory combinatory variations and usually they neglect important contrastive features of a language – features which are heard by foreigners but not by native speakers, because for native speakers such a pronunciation may seem automatic, the only possible in human speech, and as such it is not recognized. In the paper I am going to show some examples of such situations.

**The Argument from Silence? Diglossia and the identification of Sudan Arabic dialects.**

*R. Thelwall* (University of Calgary)

1. Diglossia and Arabic
2. Sudan Arabic and it’s place in Arabic dialectology
3. What do we know about Sudan Arabic dialects?
4. What do we know about neighbouring varieties and how far are they relevant?
5. A brief survey of sources
6. What do we know about Diglossia in Sudan?
7. Bilingualism and Arabic in the Sudan
8. Pidgin and creole varieties of Arabic
9. Jargons?
10. Some examples from case studies

Assertions about unique and specific dialectal origin for Sudan Arabic-speaking refugee claimants pose particular problems for contra-expertise. Published data on dialectal variation is sparse and scattered throughout the few sources. Terms such as ‘Northern
Sudan’ and ‘Central Sudan’ are employed with varying imprecision, usually to obfuscate. The notion of diglossia is difficult to use in support of the rejection of deterministic claims – e.g. “X does definitely not speak the Arabic of Y”. The status of Arabic among Sudanese for whom it may be a second or ‘parallel’ language is not adequately documented in publication and the contra-expert often has to make claims based on their own experience (and narrow competence) with little collateral. Contra-expertise often has to make claims about the absence of defining features rather than their presence (the Argument from Silence).

The linguistic constructs of phonology, morphology and syntax often provide no easy basis for distinguishing varieties given the normally wide range of performance associated with diglossia and its overlapping varieties. This paper surveys the status of our formal knowledge about Arabic in the Sudan, highlights some sample claims from ‘Determination of Origin’ cases and discusses the problems of arguing against claimed shibboleth varieties of Arabic.