

# Minimal Parameters of Linguistic Analyses for the Determination of Linguistic Background and Place of Predominant Socialization<sup>1</sup>

Lutz Rzehak (Berlin Humboldt-University)

## Introduction

“The language spoken by somebody and his or her identity as a speaker of this language are inseparable: This is surely a piece of knowledge as old as human speech itself. Language acts are acts of identity.”

Andree TABOURET-KELLER, “Language and Identity”, Florian COULMAS (ed.), *Handbook of Sociolinguistics*, Oxford: Blackwell, 1997: 315

“... it is always the individual that creates for himself the patterns of his linguistic behavior so as to resemble those of the group or groups with which from time to time he wishes to be identified, or so as to be unlike those from whom he wishes to be distinguished.”

R.B. LE PAGE, Andree TABOURET-KELLER, *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*, Cambridge: Cambridge University Press, 1985: 181

If language acts are acts of identity it must be possible to draw reasonable conclusions about a person's identity from the way one speaks. The idea of linguistic analyses for the determination of origin is based upon this axiom. Concerning possible results of linguistic analyses for the determination of origin it has been widely accepted meanwhile that they cannot be used reliably to determine national origin, nationality or citizenship because these are political categories which are not necessarily connected to language. Sometimes linguistic analyses can be used, however, to draw reasonable conclusions about a person's linguistic background and place of predominant (or secondary) socialization.

Although the general objectives of linguistic analyses within the asylum procedure are more or less clear the methods and minimal requirements of such analyses still remain in dispute. I argue that linguistic analyses for the determination of linguistic background and place of predominant socialization must be done on a scholarly basis. If linguistic analyses within the asylum procedure shall be carried out in accordance with the rule of law, the methods and the results of the analy-

---

<sup>1</sup> Reviewed version of a working paper presented at the *Workshop on Linguistic Analyses within the Asylum Procedure*, held 2008, 22-24 July, in Lausanne, Switzerland.

ses must be *provable* and *verifiable*. This can be guaranteed only with a scientific approach. Linguistic assessment of legal relevance cannot be based upon intuition or upon some shibboleths only. A wide range of factors must be considered which can influence or characterize a person's linguistic behavior. This paper is aimed to define indispensable parameters of linguistic analyses for the determination of linguistic background and place of predominant socialization.

The linguist who is appointed for such analyses must have a good knowledge not only of the language(s) in question but of the sociolinguistic situation in the claimed region(s) of origin and/or socialization as well. It is self-evident that the expert must be able to investigate and to describe distinctive features in all basic sub-areas of language structure like phonetics, morphology, syntax, and lexicon in a scholarly way. Besides that the expert must be acquainted with dialectology, with stylistic norms and rules of pragmatics, and he or she must know basic rules and determining factors of individual linguistic behavior. Needless to say the expert's knowledge must be up-to-date and he/she must know recent changes and developments in the language(s) under investigation.

Our linguistic (including sociolinguistic) knowledge of some regions of the world might be insufficient to fulfill all parameters to be outlined here. We must acknowledge the fact that in these cases a reliable analysis is impossible and cannot be conducted.

In linguistic analyses for the determination of linguistic background and place of predominant socialization the object of investigation is not the *language*, but *language in use* or the *speech style* of a person as shown in a given communication situation (usually a recorded interview). In a wider meaning the object of investigation is *linguistic behavior*. For a responsible, provable and verifiable assessment of speech style and linguistic behavior a set of criteria must be checked in a multilateral system of several coordinates as shown below. The method of such linguistic analyses can be described rather as *calibration of a person's linguistic identity* than as identification. In the final report all steps of the analysis must be described and all parameters of relevance must be assessed.

## 1 The Applicant's Statement

The phenomenon to be studied and checked with the help of linguistic analyses is a person's statement about his/her place(s) of origin and/or predominant socialization. As much as possible personal information should be asked during the interview which is made for the linguistic analysis. This can help the linguist to formulate a reasoned hypothesis about the supposable linguistic behavior and about distinctive linguistic features to be expected. Ideally the applicant's statement should contain information about:

- Region(s) of origin and/or socialization
- Native language and knowledge of other languages
- Biographical data (ethnic belonging, age, family background, social status, places and duration of residence and/or migration, education, occupation, religion etc.)

All information of that kind can be asked without revealing a person's identity (name) if this is necessary for reasons of data protection.

**In the final report all relevant personal data as presented by the applicant must be summarized.**

## 2 Hypothesis

On the basis of an applicant's personal statement the linguist must frame a hypothesis regarding the expected linguistic behavior and speech style, i.e. the linguist must make a *reasoned proposal* suggesting a possible correlation between the stated origin, socialization etc. and supposable linguistic phenomena. The hypothesis must answer the question:

- What linguistic behavior, which speech style and which other linguistic features can be expected according to a speaker's statement about his/her place(s) of origin and socialization, ethnic belonging, social status, educational level, etc?

The hypothesis is based upon the linguist's knowledge of the sociolinguistic situation in the claimed region(s) of origin and/or predominant socialization.

The characterization of the sociolinguistic situation in that region(s) should not be limited to the most obvious question which language a person with the stated origin and socialization should likely speak. Other questions to be checked when formulating the hypothesis are as follows:

- Which language(s) are generally spoken in the claimed region(s) of origin and/or socialization?
- Are there distinguishable language varieties (dialects, sociolects etc.) and of which ethnic or social groups are these varieties characteristic? Which variety can be expected with regard to the applicant's statement?
- Do these languages have different norms (classical, modern literary/formal-spoken, educated colloquial, local dialects) and in which communication situations are they commonly used? Which variety would likely be used in the interview?
- Are people in the mentioned region monolingual, bilingual or multilingual? Which patterns of multilingualism are characteristic of that region?
  - Diglossia (with a structural functional distribution of the languages/varieties involved)
  - Ambilingualism (a clearly distinguishable functional distribution is not observed)
  - Bipart-lingualism (more than one language can be used in a small area, but the majority of speakers are monolinguals)
  - Code-switching
  - Pidgins, creoles etc.

In the final report the linguist must formulate a reasoned hypothesis what linguistic behavior, which speech style and which other linguistic phenomena can be expected according to a speaker's statement about his/her place(s) of origin and socialization, ethnic belonging, social status, educational level, etc. on the one hand, and according to the linguist's knowledge of the sociolinguistic situation in these region(s) on the other hand.

### 3 Verification of the Hypothesis

The hypothesis must be evaluated and verified on the basis of linguistic data as presented in the interview which was made with the applicant for the linguistic analysis. In terms of methodology the verification of the hypothesis can be described as *calibration of a person's linguistic identity*, i.e. the linguist must study the relationship between the applicant's linguistic behavior as shown in the interview and distinctive linguistic features as expected according to the hypothesis. The method of calibration of a person's linguistic identity is based upon the idea that no two speakers are identical in linguistic production and that speech community members share essentially not only the same grammar, lexicon, stylistic repertoire, and phonological inventory, but also norms for interpretation of speech and variation.<sup>2</sup>

#### 3.1 *The Interview*

##### 3.1.1 Usability of Linguistic Data

The recorded interview must be assessed in general terms of quantity and quality regarding usability for linguistic analysis.

- Does the recorded interview contain sufficient linguistic material to demonstrate the applicant's linguistic behavior adequately?
  - How long was the interview in total?
  - How long was the speaking time of the interviewed person approximately?

In an interview with a length of 50-60 minutes on an average at least half of the total speaking time should be sufficient to assess a person's speech style and linguistic behavior. Exceptions depend on how many languages were used in the interview, what a person said and with what kind of utterances the speaking time was filled.

- Is the technical quality of the recording sufficient for a thoroughly analysis of phonetic and other features?

---

<sup>2</sup> Cf. Peter PATRICK's paper on "Sociolinguistic principles and issues of expertise in the LADO process" presented at the same *Workshop on Linguistic Analyses within the Asylum Procedure*, held 2008, 22-24 July, in Lausanne, Switzerland.

The final report must contain a statement concerning the assessment of the recorded interview in terms of quantity and quality with regard to usability for the linguistic analysis.

### 3.1.2 Communication Situation Variables

Linguistic behavior depends much on surrounding conditions. Although the interviewer usually encourages the interviewee with different methods to speak in a “most natural way”, the interview must be regarded as a quite specific communication situation where different factors can influence or even deform a person’s linguistic behavior. For the characterization of such communication situation variables the following points and questions must be checked:

#### **Language(s) and varieties**

- Which language(s) and/or language varieties were used during the interview?

#### **Linguistic competence**

- Are both the interviewer and the interviewee native speakers of the language(s) which was/were used for conversation? If not, how is the level of linguistic competence assessed?<sup>3</sup>
- Were there cases of misunderstanding? If yes, how can they be explained?

#### **Language accommodation**

- Was the speech style, probably, influenced by the formal character of the interview?

Some languages and/or varieties may be regarded inappropriate to the formal character of the interview. Other languages may not have the necessary vocabulary to express a person’s idea about a subject which was touched on during the interview. This can happen, for example, when an individual comes from a multilingual society with a structural functional distribution of the languages/varieties involved. For these and similar reasons it can happen that local dialects, colloquial norms, or creoles were not used at all during the interview although they would be the best proof of a person’s linguistic background.

- Did the interviewed person adjust her/his speech style towards the speech style of the interlocutor or vice versa?

Speech acts are acts of power. We change the way we communicate when in differing communication situations. Individuals speak to their parents differently than to their peers; no one would talk the same way to his boss as he would to friends etc. The interviewer is most likely seen as a representative of state administration. For this and similar reasons a person might try to please the interviewer by adjusting his or her speech style to

---

<sup>3</sup> Possible factors to be considered here are described below on p. 9-10.

that of the interlocutor. A speaker can more or less unconsciously suppress distinctive features of a language/variety which would help to draw reasonable conclusions about his/her linguistic background.

- Were there indications of feigning a false linguistic identity? Possible indications of feigning a false linguistic identity are among others:
  - Changes in speech speed
  - Mid-utterance pauses for thinking
  - Other speech disfluencies
  - Self-corrections
  - Hypercorrections etc.

**The final report must contain a statement saying which language(s)/varieties were used during the interview and how the command of these language(s)/varieties of both the interviewer and the interviewee is assessed. The final report must contain a statement concerning possible language accommodation as described above. It must be stated whether the speech style of the interviewed person is regarded as authentic or whether there were indications of feigning a false linguistic identity.**

## **3.2 Calibration of linguistic identity**

### **3.2.1 Distinctive Features of the Speech Style**

An all-embracing analysis of the language(s) or language varieties as they were used in the interview is not feasible both for practical and methodological reasons. Within the scope of the objectives of the linguistic analysis the linguist must check those linguistic features which are distinctive against the background of the hypothesis about supposable linguistic phenomena. In the majority of cases such distinctive features can be found in phonetics, morphology, syntax and in the lexicon as well. Distinctive features must be characterized in terms of quantity and quality. Therefore the linguist must answer the following questions:

- Which features are characteristic or non-characteristic against the background of the hypothesis in the following sub-areas of the language system?
  - Phonology
  - Morphology
  - Syntax
  - Lexicon

It is difficult to say in general how many features in phonology, morphology, syntax and lexicon must be checked and be described in the final report. This can vary from case to case. It should be as many features as necessary for the verification of the framed hypothesis. There can be exceptions when one of these sub-areas of the language system is irrelevant for

the characterization of a linguistic variety or its relevance has not yet been studied and is not described in linguistic literature. In these cases more attention must be given to all other sub-areas of the language system to make the calibration of a person's linguistic identity as sound as possible.

- Were these features observed in individual cases only or can they show an evident tendency in linguistic behavior?

Frequency of occurrence is an indispensable criterion for checking particular features against the background of the hypothesis. The more often a feature was observed in the linguistic behavior of a person the more significant it can usually be considered for the verification of the hypothesis. However, quantitative criteria must be checked in relation to how often a feature occurs in a language/variety in general. Sometimes, for example, a phoneme which is quite characteristic of a language/variety occurs in just a few words and can be observed, therefore, only once or twice (and sometimes never) during an interview.

- In which semantic context did these features occur?

Sometimes it is necessary to check the semantic context in which a particular feature occurred. Let's assume that in a person's speech style particular linguistic features (e.g. denotations of some realities) were observed which are non-characteristic of the language in the claimed region of origin, but characteristic of the language in a neighboring region. Let's assume also that this person had migrated to this neighboring region for a while and that these linguistic features occurred only when this person was talking about this migration but never when talking about the claimed region of origin. In this case these linguistic features are invalid arguments to doubt the claimed region of origin. On the other side, a person who is feigning a false linguistic identity can be caught on the mistake of using words, expressions, or more sophisticated linguistic forms in an appropriate context etc.

- Are these group-exclusive or group-preferential features?

The linguist must be aware of the fact that in some cases certain linguistic features are unlikely but not impossible against the background of the hypothesis. Group-exclusive forms may be taken for granted in one variety while they are quite obtrusive from speakers of other varieties. Group-preferential forms are distributed across different varieties, but members of one group are more likely to use the form than members of another group. In social interaction both group-exclusive and group-preferential forms may gain in symbolic significance in identifying people from a given local or social group. Corresponding stereotypes of particular regional and ethnic dialects are often caricatures and do not necessarily correspond to the actual use of the form by speakers from the particular speech community.

**In the final report the linguist must describe which phonetic, morphological, syntactical, and lexical features of the speaker's speech style are characteristic or**

non-characteristic against the background of the hypothesis. It is insufficient to list these features only. The linguist must characterize them in terms of quantity and quality as well in order to demonstrate which features show an evident tendency in linguistic behavior and which features can be taken for granted or are likely to be used in a particular variety.

### 3.2.2 Complementing Features of the Speech Style

In some cases it is useful or even necessary to keep an eye on other linguistic features as well. Here along with the stylistic repertoire and conversational features, the wide field of pragmatics can be quite significant. Pragmatics describes generally the forces in play for a given utterance and includes the study of power, gender, identity, and their interactions with individual speech acts. Pragmatics is regarded as one of the most challenging aspects for language learners to grasp. Rule-consistent or rule-inconsistent behavior in the field of pragmatics is a serious indication of a person's linguistic background and identity therefore. Unfortunately, with regard to most languages in question, pragmatics has not yet been studied and described to the same extent as phonology, morphology, syntax, or lexicon. Complementing features of that kind cannot be described in detail here because linguistic approaches in the field of pragmatics differ from language to language and from cultural background to cultural background to a high degree.

Even in those cases where a person's linguistic identity can be calibrated quite clearly on the basis of distinctive features in the sub-areas of phonology, morphology, syntax, and lexicon, checking pragmatics, the stylistic repertoire, and conversational features against the background of a speaker's claimed educational level and social status, can help to enhance the calibration of a person's linguistic identity and can make the final conclusions more credible. In the majority of cases the study of pragmatics and conversational features could at least be focused on the following questions:

- Do the usage of idiomatic expressions, filling (or nonsense) words, interjections, colloquial expressions, or religious formula and the stylistic repertoire correspond to the linguistic background, educational level and social status the speaker must have according to his/her biographical information?
- Were any other distinctive features and/or conversational irregularities observed in a person's speech style?

**In the final report complementing features of the speech style like rule-consistent or rule-inconsistent behavior in relation to pragmatics and stylistics, conversational features, and possible irregularities in the speech style must be characterized against the background of the hypothesis and in relation to the speaker's statement about his/her region(s) of origin and/or predominant socialization, educational level and social background.**



### 3.2.3 Temporal Features of the Speech Style

For the calibration of a person's linguistic identity it is necessary to take into consideration the time when a person left the speech community which he/she claims to belong to. Therefore the linguist must answer the following questions as well:

- Are there any indications that the speaker really left the mentioned region(s) of origin or socialization at the time she or he stated?
  - Did the speaker use or know the names of institutions, technical innovations etc. which had appeared only recently before that time?
  - Did the speaker use buzzwords or other ephemeral lexical items which would be characteristic of a certain period of time?

**The final report must contain a statement whether there were any indications that the speaker really left the claimed region(s) of origin and/or predominant socialization at the time she/he stated.**

### 3.2.4 Phenomena of Multilingualism and Language Contact

Multilingual speakers outnumber monolingual speakers in the world's population. Accordingly linguistic identity of most individuals is formed by more than one language. Multilingualism often follows more or less traceable and describable patterns which are typical of a particular region. The term multilingualism can refer to an occurrence regarding either a community of speakers where two or more languages are used or an individual speaker who uses two or more languages. Not all members of a multilingual community are necessarily multilingual; and there are different types of bi- or multilingual persons.

Features related to multilingualism and language contact must be checked against the background of the hypothesis in order to calibrate a person's linguistic identity thoroughly. Here among others the following questions must be answered:

- Is the speaker mono-, bi- or multilingual?
- If bilingual is the speaker compound, coordinate, subordinate bilingual, or something else between these types?

Compound bilinguals are fluent in both languages and for them words and phrases in both languages are in the same concepts. For coordinate bilinguals one language is more dominant than the other. The sub-group of sub-ordinate bilinguals is typical of beginning second language learners. In reality it is often difficult to assign a speaker clearly to one of these types. A qualitative characterization of a bi- or multilingual person's language command and linguistic behavior should rather be given therefore.

- Which language is the first-language ("mother tongue")?

In some cases a person's first-language can be a local dialect or a creole language which is recognized neither *de jure* nor *de facto*. Assessment

should not be based upon a bureaucratic nomenclature of recognized languages and their official names therefore, but rather on the speaker's own statement and the linguist's observation of language competence and linguistic behavior. A language which a person calls "mother tongue" must not necessarily be identical to this person's first-language, i.e. to the language which he/she speaks best and which is used in most communication situations because it is more suitable to fulfill most communication requirements. Sometimes a speech community uses a creole language which has not yet created a name of its own and which would hardly be named as "mother tongue" therefore.

- How can the speaker's command of the other language(s) be characterized?

Together with descriptive assessments like "He/she is fluent", "He/she speaks without any accent" convenient criteria among others can be the range of the vocabulary used by a speaker, and the usage of rule-consistent or rule-inconsistent grammatical forms. Sometimes it can be assessed whether a language was learned in everyday communication mainly whether it was learned more or less theoretically at school etc.

- Were cases of code-switching or phenomena like pidgins and creoles observed?

Code-switching ('swapping' between languages) as well as pidgins and creoles, are very typical phenomena in multilingual societies. In the interview cases of code-switching can most likely be expected when the interviewer belongs to the same speech community as the interviewee and when both have more or less equal command of all languages/varieties in question. Pidgins and Creoles follow own rules and can be expected only if both the interviewer and the interviewee know the norms of their usage. A person who is used to code-switching or a person who usually speaks a creole language may have difficulties to express his/her ideas only in one of the languages involved if the interviewer does not have sufficient knowledge of all codes in questions, i.e. if the interviewer knows only one of the languages/varieties which are usually used simultaneously or interchangeably within a speech community.

- Which linguistic features can be interpreted as influence of other languages on the speech style?

Often even monolingual individuals show features of linguistic interference which can be interpreted as influence of other languages on their speech style. Cases of linguistic interference can happen when a person grew up or lived in a multilingual society or when a person due to migration(s) lived in a foreign speech community for a while, etc. The whole range of code-copying (influence on phonetics, any sort of borrowings, etc.) can be checked in that regard.

**In the final report all features concerning bi- or multilingualism and language contact must be characterized in relation to biographical information and to the sociolinguistic knowledge of the claimed region(s) of origin and/or predominant**

socialization, i.e. against the background of the framed hypothesis about supposed linguistic phenomena.

## 4 Final Conclusions

On the basis of all features and parameters in speech style and linguistic behavior which were checked for the calibration of a person's linguistic identity, the linguist must draw final conclusions in relation to the framed hypothesis. Final conclusions cannot be drawn on the basis of selective features only even if they are regarded a shibboleth of a particular language/variety. Ideally final conclusions are based on the ensemble of all checked features and parameters.

Questions concerning the speech community which a speaker can be assigned to and questions concerning the probable region(s) of origin and/or predominant socialization must be answered separately because both answers are not necessarily connected to each other.

Questions to be answered in order to formulate final conclusions are as follows:

- Was the hypothesis about the supposed linguistic competence, speech style, and linguistic behavior confirmed or disproved?
  - Can the speaker generally be assigned to the supposed speech community?
  - If not which speech community can the speaker be assigned to or how can the speaker's linguistic identity be characterized?
  - Are there any linguistic indications of longer stays outside that speech community?
  - Do speech style and linguistic behavior correspond to the stated social status and level of education?
- Which conclusions can be drawn about the probable region(s) of origin and/or primary/secondary socialization?
- Are there evident sources of trouble or error? If yes they must be mentioned in the report.

**In the final report the linguist must draw final conclusions concerning the relationship between the applicant's speech style and linguistic behavior as shown in the interview, and distinctive linguistic features as expected according to the hypothesis. Ideally final conclusions are based on the ensemble of all checked features and parameters. Conclusions concerning linguistic identity and conclusions concerning the probable region of predominant (and secondary) socialization must be drawn separately.**

## Closing Words

Within the asylum procedure it seems impossible to prove afterwards whether the result of an analysis indicates a speaker's real linguistic background and place of predominant socialization. The only way to examine the probable accuracy of

linguistic analyses carried out by a particular linguist for a given linguistic area would be a series of tests with speakers who are not involved into an asylum procedure. Such tests are time-consuming, expensive, and often not feasible for practical reasons. Besides that, the idea of using empirical tests in order to examine the significance of analyses of such complex and only partially predictable phenomena like linguistic behavior seems to be naïve.

The best quality assurance is a proof that linguistic analyses comply with established requirements concerning the methods and documentation of the analysis. The minimal parameters of linguistic analyses for the determination of linguistic background and place of predominant socialization as they were outlined here are thought as a proposal to establish such indispensable requirements. Needless to say these parameters and the underlying method of calibration of a person's linguistic identity must be improved and enhanced in practice and in terms of theory of language and linguistic behavior.

© Lutz Rzehak 2008

Contact Information:

[lutz.rzehak@web.de](mailto:lutz.rzehak@web.de) | <http://www2.hu-berlin.de/zentralasien/?section=rzehak>

Humboldt-Universität zu Berlin | Zentralasien-Seminar | Unter den Linden 6 | 10099 Berlin | Germany