Research into Multilingualism in Estonia

Anna Verschik
Tallinn University, Tallinn, Estonia

The present paper concentrates on several issues relevant to research into multilingualism in Estonia. It is argued that a macrosociolinguistic approach is insufficient when not counterbalanced with microsociolinguistic studies (case studies of actual linguistic behaviour, linguistic creativity, mechanisms and practices of multilingual communication, construction of immediate and long-term strategies of multilingual communication, etc). Furthermore, it is demonstrated why the census of 2000 in Estonia gives a rather distorted picture of multilingualism. Another problem is the lack of scholarly attention to the so-called third languages (non-Russian and non-Estonian). The example of the Jewish community in Estonia shows how a minority with a supposedly common ethnohistorical background is in fact split into two separate groups (indigenous and nonindigenous) with different sociolinguistic profiles. It remains to be seen for which minority groups in Estonia the same kind of division is relevant. Finally, topics for future research (language change and identity construction, dynamics of multilingual conversation, emerging ethnolects) are outlined.

Keywords: Estonia, indigenous minorities, microsociolinguistics, multilingualism

Introduction

According to the recent census, there are 142 ethnic groups living in Estonia (http://www.stat.ee/files/eva2003/RV200102.pdf, table 7). While these groups do not obviously have an equal sociodemographic and/or cultural significance in the sociolinguistic profile of Estonia, systematic sociolinguistic studies of communities other than Russians and Russian speakers would be extremely necessary. Unfortunately, during the years after the regaining of independence, a strong tendency to exclude from (socio)linguistic investigation other speech communities has become apparent. However, even the situation of Russian in Estonia still needs thorough investigation going beyond general sociological studies and mere registration of lexical borrowing (see details in Verschik, 2004–2005).

The present paper is organised as follows: first, the importance of microsociolinguistic methodology in contact situation research will be emphasised. Second, some reasons for such a poor state of multilingualism research in Estonia will be explained. Third, the significance of linguistic creativity and possibilities of identity construction will be discussed, and finally, some lines of inquiry in multilingualism research will be suggested.

Advocating a Microsociolinguistic Approach

The relative ‘invisibility’ of the so-called third languages/speech communities is perfectly understandable in the context of Estonian sociocultural and
demographic history. Although non-Russian and non-Estonian ethnic groups are marginal in a demographical and statistical sense, they should not remain in the shadow. As Stolz (2001) has convincingly demonstrated, a long-lasting disregard towards minor languages has led to a distorted (socio)linguistic picture in mainstream linguistics. The case made by Stolz is rather clear and there is no need to discuss it further.

Unfortunately, the so-called third languages still remain ignored. For instance, two projects on accent research (2001–2003, carried out by Tallinn Pedagogical University, and 2002–2003, carried out by the latter in cooperation with Estonian Radio and the Institute of Cybernetics) have dealt exclusively with the Russian accent in Estonian. The main focus is on the implications of Russian-accented Estonian speech for language technology. So far only two studies that deal with other aspects of Russian accent can be mentioned. Rannut (2000) describes the main features of the Russian accent in Estonian and gives some practical recommendations to teachers of Estonian as L2. The second study is a BA thesis by Sapelson (2003). His sociolinguistic research uses a Matched Guise technique in order to find out how young Estonian speakers from Tallinn feel about Estonian with an accent. The target group was asked to estimate different voices: Standard Estonian, and Estonian with an American English, Russian and Finnish accent. Sapelson (2003) found that Standard Estonian was the most highly valued and that Russian-accented Estonian was valued more highly than English- and Finnish-accented Estonian. According to him, young Estonians appreciate the effort of Russian speakers who have mastered Estonian. Men tended to value female voices more highly, and women valued male voices more, but this is not for sociocultural reasons.

Unfortunately, Sapelson’s (2003) remains the only study of this kind. Apart from this, Estonian with an English and a Finnish accent has not been treated at all, let alone more ‘exotic’ accents. There is yet another aspect of accent research: recently, foreign-accented Estonian started to be used in commercials, for instance, Estonian with a heavy Swedish accent in a recent TV commercial of a pension fund provided by Ühispank, one of the biggest banks in Estonia. I believe that, although a Swedish or, say, a Lithuanian accent in Estonian is not a frequent phenomenon and may seem even exotic in the present context, such rarely observed phenomena too deserve scholarly attention.

The unsatisfactory state of research into small speech communities in Estonia is partly due to the fact that this country itself is tiny, and experts in languages such as Tatar, Armenian or Romani are lacking. However, a more significant reason lies elsewhere, namely in the prevalence of macrosociolinguistic and quantitative research. Important and useful tools as they are, large-scale macrosociolinguistic studies alone fail to give an adequate account of a sociolinguistic situation when not supported by microsociolinguistic research.

The recent census of 2000 is a remarkable illustration of the fact. First, the questionnaire contains too few questions concerning languages. Second, it presupposes that a person has only one mother tongue. Although any language can be indicated as one’s mother tongue (Question 13), all languages other than the mother tongue would automatically be placed under the heading ‘other languages that you know’ (Question 14). However, the list of languages falling
into the latter category is limited, and the options are as follows: Estonian, Russian, English, German, French, Finnish, Swedish, Latvian, one’s own national language, other languages, do not know any other languages. Remarkably, the entries ‘own national language’ and ‘other languages’ cannot be specified. It also needs to be noted that language proficiency is not defined.

All this leaves us with a set of complex problems. It is well known that the notion of mother tongue in a multilingual individual is often misleading and highly dependent on mother tongue definition criteria, such as origin, competence, function, and internal and external identification (Skutnabb-Kangas, 2000). Therefore, a multiple linguistic identity essential for many multilinguals is not allowed for in the census. There is more to it: in an internally conflicted speech/ethnic community, the concept of ‘own national language’ is often vague and arguable. Jews are a classical example of an internally conflicted community (see Verschik, 2000, 2002 for a discussion of Jewish multiple identities in Estonia). In such cases, the census data are to be interpreted with caution.

It is a well known fact that the way in which a question is formulated affects its answer. If, instead of mother tongue, we settle for ‘most frequently used language’ or ‘home language’, there is no guarantee that we would finally get an undistorted picture of multilingualism. Frequency of use definitely provides important information, however, in the case of small and/or dispersed groups, this information has only an indirect connection with self-identification and preferences in language use. It may be impossible to name a ‘home language’ in a situation of stable and complicated multilingualism (see Clyne, 2003: 21ff for various difficulties connected with census questionnaires).

Thus, alongside large-scaled macrosociolinguistic research, other tools are also needed, especially for research into small speech communities. Linguistic biographies are but one example of possible techniques (Franceschini, 2002a). It is an empirical term designating a methodological approach to investigate the dynamics of language choice, linguistic preferences and competence in a multilingual individual (Verschik, 2002: 39).

Traditional research that is preoccupied mainly with lexical borrowing and, to a lesser extent, reciprocal impacts in phonology and morphology, lies on the other end of the continuum. For instance, a focus on lexical borrowing and an almost total disregard of sociolinguistic factors is typical of earlier studies of Estonian influence on Russian dialects in the Lake Peipus region (Heiter, 1970, 1975; Mürkhein, 1970) as well as of contemporary research on these dialects (Burdakova, 1997; Burdakova & Burdakova, 2000) and of modern (Standard) Russian in Estonia (Külmöja, 1999). The aforementioned studies certainly contain valuable data; however, they are mostly descriptive and often fail to identify contact phenomena (codeswitching and its relation to lexical borrowing, language shift versus language maintenance etc.).

Although it has been demonstrated by various scholars that structural properties alone do not determine the result of language contacts, the view that it is the structure that sets constraints on possible outcomes and defines what happens remains popular (Thomason, 1997; Thomason & Kaufman, 1988 and Bilingualism: Language and Cognition 1999, 2/2 for a detailed discussion of the relevance of sociolinguistic history especially Sebba, 1999; Singh, 1999).
Thomason (1997, 2000) shows that speakers’ attitudes, their linguistic creativity, and change by deliberate decision are crucial to language contact results. All these factors affect the structural properties and in the long run may influence the sociolinguistic profile of a given speech community.

It seems that a microsociolinguistic approach to multilingualism is needed at this stage. In studies conducted by Auer (1998: 3–4; 1999), a microsociolinguistic analysis has proved to be a promising tool in its application to codeswitching research: this approach leaves space for (individual) linguistic creativity/changes by deliberate decision. In codeswitching research this approach helps to fill the gap between a strictly grammatical and macrosociolinguistic approach (Meeuwis & Blommaert, 1994). I believe that a microsociolinguistic approach might be helpful with research into multilingualism in general. In the next section we shall see how a focus on linguistic creativity can contribute to a general picture of multilingualism. The division of labour between traditional, microsociolinguistic and macrosociolinguistic approaches to multilingualism research is pictured in Table 1.1

At present, more data collection and mapping of language data are needed. I believe that the application of qualitative microsociolinguistic methods to research into multilingualism could prove productive.

**Change by Deliberate Decision (Linguistic Creativity) and Identity Construction**

An image of a language as a well defined, clearly delimited system has been essential in linguistics for a long time and still remains attractive to many
researchers. Nevertheless, recent developments in contact linguistics and multilingualism research point to another direction. Language is neither a 'fortress' (Muysken, 2000) nor a 'monolith' (Backus, 1999). Therefore, behind seemingly simple and clear signs such as 'Russian', 'Estonian', 'English' etc. there stands not one variety but a whole range of them. In the same vein, research into bilingualism and second language acquisition have made it explicit that a multilingual person is not merely a sum of several monolinguals but that s/he can function in linguistic modes that are unavailable to monolinguals (Grosjean, 1995, 2000).

A multilingual grammar is not merely confined to new combinability rules of grammatical and lexical material belonging to several monolingual grammars. Structural convergence is a striking feature of multilingual speech, as well as various compromise forms (Clyne, 1987, 1997). Both convergence and compromise can be a result of deliberate decisions made by speakers. According to Thomason (1997), even a passive knowledge of another variety may lead to changes in the first language. In any study of multilingualism the following question should be asked: are there monolingual speakers of the variety in question? Trivial as it may appear on the surface, the answer to this question has crucial sociolinguistic and structural consequences. For instance, a study of Finnish nouns in the Russian speech of the indigenous Russian minority in Finland (Leisio, 2001) demonstrates that in the generation of Russian speakers who are all bilingual, there is a clear tendency not to integrate Finnish nouns into the Russian matrix. It is not relevant whether a Finnish noun fits into Russian noun declension classes, because sociolinguistic factors in this case outweigh structural ones.

A change by deliberate decision may result in new patterns, forms and structures that differ from results of 'natural' (i.e. nondeliberate) contact-induced change. For instance, Golovko (1994) makes this point by drawing a border between ordinary codeswitching and codeswitching games (i.e. deliberate use of codeswitching for satirical or comical purposes). He claims that the two types of codeswitching may evoke different structural properties. A detailed analysis of codeswitching data between Estonian and other co-territorial languages, as well as discussion of relevance of Estonian-X codeswitching data to the general theory of codeswitching remains beyond the scope of the present article (see Verschik, 2004–2005, on Russian–Estonian codeswitching). Instead, some examples from empirical evidence in the speech of an Estonian 12-year-old girl studying English and Russian at school will be chosen to confirm the above claim (Golovko, 1994).

She experiments with a combination of three languages in the same utterance. There is a strict rule of the distribution of negative particles ne and net(u) in Russian: the particle ne 'not' negates a particular word (ja ne znaju 'I do not know', doroga ne dlina 'the way is not long' etc), while net(u) 'no, (there is) none' negates a noun in an existential sentence (u menja net(u) knigi 'I have no book') or may be an answer to yes-or-no question (Ty znajesh? 'do you know' – Net 'no (I don’t)'). Netu is used only for the negation of nouns in existential sentences; is not accepted in the written language and in serious registers and is considered to be colloquial. Although the girl is aware of the distribution rule and uses the negative particles correctly when speaking
Russian, she has developed a game of (mixed) register with generalised *net(u)*. Consider (1) where Estonian is in italics, Russian is bold and English is underlined:

(1) \[
\text{Raamat} \quad \text{net} \quad \text{interesting},
\]
\[
\text{book} \quad \text{not} \quad \text{interesting}
\]

‘the book is not interesting’

Cf. with Standard Russian in (2):

(2) \[
\text{Kniga} \quad \text{ne} \quad \text{interesnaja}.
\]
\[
\text{book} \quad \text{not} \quad \text{interesting}
\]

‘the book is not interesting’

First of all, it is possible that the child treats English *not* and Russian *net* as homophonous diamorphs and consciously creates ambiguity. Secondly, codeswitching between three languages in the same utterance does not happen as frequently as between two languages, at least in Estonia. Thirdly, there is no evidence that the generalised *net* is used in ‘real’, nondeliberate Russian–Estonian codeswitching (i.e. as opposed to the codeswitching game developed by the child). To the best of my knowledge, the use of the negative particles in sentences with Russian–Estonian codeswitching occurs in accordance with the rules of the Russian grammar, as in (3). Example (3) is taken from a bilingual TV show. A balanced Russian–Estonian bilingual journalist explains his opinion. The negative particle before the Estonian noun *kättemaks* ‘revenge’ is *ne*, as it is expected according to the above-mentioned rule. The Russian items are in bold:

(3) \[
\text{Nu, éto, ja by skazal, ne kättemaks}
\]
\[
\text{so this I would say not revenge}
\]

‘so, I would not call it revenge’

Compare this to Standard Russian in (4), with the same meaning:

(4) \[
\text{Nu, éto, ja by skazal, ne mest’}
\]
\[
\text{so this I would say not revenge}
\]

‘so, I would not call it revenge’

Thus, the utterance (1) with the playful deliberate use of the generalised *net(u)* is different both from nondeliberate codeswitched utterance (3) and from monolingual Russian (2) and (4).

Linguistic creativity is not limited to grammatical aspects. It can also become a striking feature in phonology. Pronunciation of certain English words with an Estonian accent, which, among other things, means the application of the so-called Estonian third quantity (over-long as opposed to short and long) marked with (‘), can make these words meaningful in Estonian, allowing thus for a bilingual pun: *some* becomes *samm* ‘step’. The English items are underlined, the Estonian ones are given in italics.
I want some komm.

*I want some sweets. I want some/step sweets*

(she utilises the new meaning) Samm komm!

This instance neatly fits into the picture of bilingual creativity as given by Hinnenkamp (2003). Referring to ‘bilingual language players’ (Hinnenkamp, 2003: 31), he stresses that they are highly aware of language resources and the potentials thereof, often exploiting their bilingualism for boundary crossing by fusing and blending words. In Estonian, linguistic creativity can be observed also on a societal level. Linguistic creativity and ‘negotiation’ in the terms of Thomason (1997) play a significant role in the communication between Estonian salespersons, shopkeepers, bus drivers, hairdressers etc. and their Finnish clientele in Tallinn. To the best of my knowledge, no research of this kind of communication has ever been undertaken. However, such a study could be promising, as a whole range of in-between varieties exists, whereas the borders between the varieties are not clearcut. It is usually thought that it is an Estonian speaker who creates a compromise variety and makes an effort in the direction of Finnish. While this may be true in terms of frequency (i.e. Estonians speak Finnish or quasi-Finnish rather than vice versa), the other option should not be disregarded. Observe utterance (6) by a Finnish customer who is trying to express her wish to a salesperson in Estonian and Standard Estonian in (7):

(6) Üksi [the ending is hardly audible] hapukoori.
   ‘one sour-cream’

(7) Üks hapukoor.
   ‘one sour-cream’

Historically, Estonian has lost the last vowel of the stem in the nominative singular, while Finnish has preserved it. Note that in Finnish ‘sour-cream’ is smetana (borrowed from Russian); also, hapukoori ‘sour-cream’ in (6) is Estonian nominative singular with an added vowel that imitates the usual state of affairs where Finnish has the last vowel of the stem (cf. Estonian pool ‘side’ and Finnish puoli ‘ibid.’, Estonian üks ‘one’ and Finnish yksi ‘ibid.’). The result is hapukoori, which may be analysed as a compromise form: an Estonian stem that has no cognate in related Finnish, but, at the same time, the noun ‘looks like’ Finnish nouns usually do.

As Le Page and Tabouret-Keller (1985) have made explicit, any utterance is an act of identity. Utterances made by L2 learners may also be viewed in the same light (Pavlenko, 2000: 285). I believe it would be reasonable to distinguish between an immediate identity (created or maintained for a particular conversation, interaction, situation etc.) and a fundamental identity that pursues long-term goals (being a member of one or several speech communities, of an ethnic and/or religious group etc.). Note that concealing the knowledge of a certain language may also be a means of identity creation. This has been relevant in Estonia during the last decades of Soviet domination when the numbers of Estonians that claimed to know Russian were
surprisingly low in the censuses of 1979 and 1989. On an individual scale, many Estonian speakers would avoid speaking Russian when possible. This type of linguistic behaviour was essential in a circumstance of conflicting language hierarchies: the official Soviet one and that of Estonian speakers.

The division between indigenous and nonindigenous minorities/speech communities in Estonia is relevant at least in the case of Russians (Berg, 1999) and Jews (Verschik, 2000). On the basis of criteria such as a long-established relationship with the country of residence, citizenship, a sufficiently representative number of speakers, a clearly discernible cultural, linguistic, ethnic and/or religious character, and motivation to preserve one’s own ethnolinguistic identity, Russians, Germans, Swedes, Latvians and Jews were acknowledged as indigenous minorities in pre-1940 Estonia. Tatars and Roma on the other hand did not qualify because of their small number of speakers (less than 3000). After the regaining of independence in 1991, all pre-war minorities are regarded as indigenous, regardless of their number. This adds Tatars and Roma to the previous list (see details in Viikberg, 2000). It has to be emphasised that the Soviet-time newcomers (mostly Russian monolinguals) are not usually considered as a minority (see details in Rannut, 2004). What is sometimes called ‘Russian-speaking population’ is, in fact, a culturally and socially heterogeneous group that is becoming more and more differentiated (Vihalemm, 2002). Self-identification, linguistic preferences and value systems can differ to such an extent that the members of an indigenous minority may have more common features with Estonians than with nonindigenous members of the same ethnic groups.

The example below describes a Jewish middle-aged woman who belongs to the indigenous minority (called by its members Estonian/Baltic Jews as opposed to the Soviet-time newcomers of Jewish origin who are mostly Russian monolinguals). She acquired German and Estonian in her childhood, Yiddish in her adolescence and, to some extent, Russian after WWII. Although she uses Russian when needed, she preferred to conceal her knowledge of Russian whilst visiting Israel. Apart from friends and relatives it was relatively hard to find speakers of German or Yiddish (let alone Estonian speakers). Thus, she sometimes preferred remaining without any means of communication than being identified by outsiders as a Russian monolingual: ‘It is hard to explain to the general public in Israel what Estonia is. Besides, I did not like the idea that Estonia would be thought of as a province of Russia. Neither did I want to be identified with those Russian Jews who do not want to learn Hebrew. I rather decided to remain silent.’

This example demonstrates how the issue of indigenousness is relevant in Estonia and how proficient multilinguals may choose to avoid using Russian because they are afraid to be associated with Russian monolinguals. However, we know almost nothing about linguistic behaviour and identity construction among the members of indigenous minorities.

Ethnolects and their Further Development

An ethnolect is a variety that marks speakers as members of ethnic groups who originally used another language or distinctive variety (Clyne, 2000: 86).
Ethnolects are often regarded in the context of (massive) imperfect L2 acquisition (L1 interference). However, this is not the only way for ethnolects to emerge. Clyne (2003: 157) enumerates several other possibilities: matrix language turnover, pidginisation, semantic transference and specific prosodic features. Another option can be added to these: new combinability rules of L2 morphemes.5

It is useful to distinguish between ethnolects that mark a single ethnic group and those that emerge among various ethnic groups, called multi-ethnolects (Clyne, 2000: 86–97). The so-called Rinkeby Swedish, a variety of Swedish spoken among immigrants in a suburb of Stockholm (Kostinas, 1998) is an example of the latter.

Ethnolects have not been studied in Estonia; however, in a situation of large-scaled L2 acquisition at least the possibility of ethnolect emergence should be considered. According to the 2000 Census (inaccurate as it may be), 44% of Russian speakers claim to know Estonian (as opposed to 15% in 1989). Various sociologists have found that differentiation within the Russian-speaking population is becoming even more pronounced as far as proficiency in Estonian and the symbolic importance of the language are concerned (see references in Vihalemm, 2002). School statistics demonstrate that the number of students in Russian-medium schools steadily decreases every year by 4–5%, whereas enrolment in Estonian-medium education establishments increases (Rannut, 2004). Therefore, there is a potential group of speakers shifting from Russian and to Estonian, and ethnolect emergence should be considered as a possible future scenario.

There exists a solid body of literature on the dynamics of ethnolects (Androutsopoulos, 2001, 2003; Fishman, 1985; Jacobs, 1996; Kostinas, 1998). An ethnolect may turn into a special register in the next generation of speakers who speak the full version of a target language. On the other hand, ethnolects or some ethnolectal features find their way into mainstream language use via contact between the speakers. The depiction of ethnolectal speech in media, fiction, popular culture etc. can lead to the creation of stereotypes (not necessarily negative or racist). From the media, ethnolectal features may diffuse into the general use (Androutsopoulos, 2001).

A careful study of old and new Estonian-language fiction as well as of modern media (commercial, TV serials etc.) could contribute to a general understanding of the dynamics of ethnolects. For instance, the Estonian of (Baltic) Germans as depicted in Estonian fiction would be a good candidate for this sort of study. One has to bear in mind that the cover-term ‘Estonian as spoken by Russians’ may in fact designate several distinct varieties in different places used by speakers with various degrees of proficiency in Estonian.

**Conclusion**

Apparently, the present sociolinguistic setting in Estonia provides most interesting examples of emerging changes in linguistic behaviour, new identities and new contact situations. The poor state of research into multilingualism in Estonia can be explained by the following two factors: the lack of a sufficient body of microsociolinguistic studies and an ignorance of the
so-called ‘third’ languages and communities. At this point, data collection and qualitative analysis are needed.

Taking into account the aforementioned aspects of the sociolinguistic situation, future research into multilingualism in Estonia could concentrate on the following topics:

1. Language change and identity construction (by ‘negotiation’, compromise and other means). This topic is closely connected with 2 below.
2. The dynamics of bilingual conversation. Solid qualitative case studies of multilingual communication (for instance, such as interaction in the market place; cases of different communication strategies used by Russian speakers who have successfully acquired Estonian; convergence and spread of convergent and compromise forms in the monolingual use etc.) would bring new facts to the attention of scholars in the field.
3. Differences in the sociolinguistic profile between indigenous and non-indigenous communities. The case of indigenous versus nonindigenous Jews clearly demonstrates the relevance of such a differentiated approach. The principle characteristics may be described in terms of a multilingual indigenous community versus a Russian monolingual nonindigenous community. It is clear though that other communities showing the above-mentioned internal division may exhibit other important sociolinguistic characteristics in addition to multilingualism versus monolingualism.
4. Ethnolects and their dynamics. The logic of the present contact situation (i.e. differentiation within the Russian-speaking community as far as acquisition of Estonian is concerned) may lead to the emergence of ethnolects. While the general theory of ethnolects is very recent, evidence from Estonia might contribute to the general understanding of ethnolect typology, formation and development.

Correspondence

Any correspondence should be directed to Assistant Professor Anna Verschik, Tallinn University, Narva Road 25, 10120 Tallinn, Estonia (anna.verschik@tlu.ee).

Notes

1. Of course, other aspects of and approaches to multilingualism exist: cognitive, psycholinguistic etc. These are not to be considered in the present article.
2. Rampton (1999) stresses that not only L2 but also ‘foreign languages’ (studied at school as a subject) are relevant for research into multilingualism.
3. Franceschini (2002b) uses the term quasi-Italianisch (quasi-Italian) to designate a variety of Italian spoken by German-speaking salespersons in Switzerland. It is implied that the speakers are aware that their variety is not an attempt to approximate a ‘full version’ of Italian but rather a compromise.
4. The terms ‘immediate identity’ and ‘fundamental identity’ are empirical shorthand concepts to distinguish between two pragmatically and sociolinguistically different situations in identity creation/maintenance.
5. This is the case in certain Jewish ethnolects (Jewish Russian, Jewish American English etc.) where word stems and derivational suffixes come from a non-Jewish language but their combinability rules are unique to Jewish ethnolects (for references see Gold, 1985; Verschik, 2003).
References


