Lívia Körtvélyessy

Productivity and Creativity in Word-Formation
A Sociolinguistics Perspective

Abstract

The paper deals with a sociolinguistic approach to productivity and creativity in word-formation. It presents research carried out to find a link between the user of a language and the language as a system; the research draws on Horecký's (2000) observation of a lack of attention paid to the relation between a language and a society, between a language as a system and language users. The paper focuses on sociolinguistic factors of gender, age, education, occupation, and language background, and their influence on productivity in word-formation in two groups of bilingual speakers (Hungarian-English and Hungarian-Slovak). The focal part of the paper is an analysis of the data gained through the questionnaire – correlations between productivity and the specific sociolinguistic factors are evaluated, with special emphasis on the correlation between productivity and language background because it turned out to be an independent and autonomous sociolinguistic factor. In general, the research has confirmed the hypothesis of the influence of sociolinguistic factors upon the naming strategies, while the strongest influence was observed for age and occupation.

1. Introduction

Productivity, one of the universal properties of language, manifests itself in word-formation whenever a speech community needs to give a name to an object of extra-linguistic reality. Productivity has become one of the central issues in research into word-formation (for example, Bauer 1983, 2001, Kastovsky 1986, Plag 1999, Baayen 1989, 1992, 1993, Baayen and Lieber 1991), and the same applies to linguistic factors which affect/restrict the productivity of word-formation rules (for example, van Marle 1986, Fabb 1988, Rainer 1993, 2005). Strangely, there has been hardly any discussion on extra-linguistic (sociolinguistic) factors influencing the productivity in word-formation. The only exception appears to be Štekauer et al. (2005). This paper examines the role of language background in the naming process. The paper provides a theoretical framework of the research (sections 2 and 3), presents an experimental research (section 4), and analyzes and comments on the research results (section 5).

2. Theoretical Framework

The hypothesis central to our research was that new complex words result from an interplay between sociolinguistic factors (the creative aspect of word-formation) and the pressure that word-formation rules impose on individual word-formation strategies (the productive aspect of word-formation). In other words, a particular object of extra-linguistic reality can usually be approached by various naming strategies the selection of which is determined by their respective productivity and also by the influence of one's naming preferences. Our informants were two groups of bilingual speakers – Hungarian-Slovak and Hungarian-English bilinguals. Our research was based on the following theoretical principles:
- an onomasiological theory of word-formation (Štekauer 1998, 2005)
- a theory of creativity within productivity constraints as developed by Štekauer, who maintains that, "[i]t is the interaction between the conceptual, the onomasiological, and the onomatological levels that – within the limits of productive types and rules and the relevant constraints – provides certain space for a *creative* approach to word-formation" (Štekauer

et al. 2005: 224)
- a concept of bilingualism as a social phenomenon, resulting from the interrelation between language and culture

3. Onomasiological Theory of Word-Formation

Since our analysis of the research data is based on an onomasiological approach to word-formation, this section briefly outlines its basic principles.

Horecký (1983: 19) maintains that any act of word-formation may be represented in the following way:
(1)

LEVEL	UNITS
1. Extra-linguistic reality	Objects
2. Intellectual (Logical)	Logical predicates
3. Semantic	Semantic components
4. Onomasiological	Morphemes, words
5. Onomatological	Affixes, words
6. Phonological	Morphemes, phonemes

According to Dokulil (1962), the onomasiological level offers different options for the structuring of the object to be named, in view of its expression in the given language. In principle, an onomasiological structure consists of two elements. The phenomenon to be named is first classed with a certain conceptual group and functions as *onomasiological base*. Then, within the limits of this group, it is determined by an *onomasiological mark*. For example, the onomasiological base of *novelist* is Agent, the onomasiological mark is Result (of Action). Importantly, the mark can be subdivided into the determining and the determined constituents. As extensively discussed in Štekauer (2005b) the determined constituent of the mark is reserved for ACTION which may be regarded as a crucial element for a correct interpretation of the relation between the base and the determining constituent of the mark.

While the onomasiological level establishes a cognitive framework for the act of naming its individual categories may but needn't be expressed by morphemes retrieved from the Lexicon at the onomatological level. This gives rise to five basic onomasiological (naming) types:

OT1: all three constituents of the onomasiological structure are expressed by morphemes at the onomatological level:

(2)   Result   –   Action   –   Agent
        novel   *write*   *er*

OT2: the determining constituent of the onomasiological mark is not expressed morphematically at the onomatological level:

(3)   Result   –   Action   –   Agent
                         write   *er*

OT3: the determinined constituent of the onomasiological mark is not expressed morphematically at the onomatological level:

(4)   Result   –   Action   –   Agent
        novel   *ist*

OT4 the onomasiological mark cannot be analysed into the determining and the determined constituents:

(5)   Negation   –   Quality
        un   *happy*

OT5 corresponds to what has been traditionally labelled as conversion or zero-derivation.

This approach establishes a framework for an onomasiological approach to productivity (Štekauer 2005). Productivity of onomasiological types is related to a particular cognitive category (Agent, Patient, Instrument, Quality, Action, Location, Result, Object, etc.). For each cognitive category, there is a universal tendency in a particular language to prefer one of the five onomasiological types which, however, usually does not preclude the other types from being employed. This gives a considerable space for a language user's creative approach to the naming act. Then, the productivity of onomasiological types is calculated as the proportion of the individual onomasiological types of all complex words falling within a particular cognitive category.

In addition to the productivity at the level of onomasiological types, productivity can analogically be calculated for:
- **word-formation types** (such as [Object ← Action – Agent]; [Action – Agent]; [Location – Action – Agent]; [Result ← Action – Agent]; [Instrument – Action – Agent]; [Manner – Action – Agent] for the cognitive category of Agents);
- **morphological types** (such as [N+V+er] as in *wood-cutter*; [N+ist] as in *novelist*; [V+er] as in *writer*; [V–>N] as in *cheat*; [N+s+man] as in *oarsman*; [A+N+ian] as in *transformational grammarian*; [N+N] as in *bodyguard*, etc.)

4. Research Description

4.1. Sample of Informants

The aim of the research was to analyse the influence of language background on the coining of new complex words. The data for two typologically different languages – Slovak and English – were obtained by means of a questionnaire (see the Appendix). The target groups of our research were bilingual Hungarian-English and Hungarian-Slovak speakers who had acquired both languages in natural environment from native speakers who used both languages for everyday communication. The language shared by both groups of informants was the Hungarian language.

The questionnaires were distributed via e-mail and surface mail communication as well as through personal contact. The most successful way of how to contact Hungarians living in English speaking countries turned out to be visits of Hungarian chatrooms on the Internet. Altogether 328 questionnaires were returned. Out of them, 146 English and 142 Slovak questionnaires were suitable for the subsequent analysis, amounting to a corpus of 1252 English and 1195 Slovak complex words.

For the purpose of our research, the following factors were taken into consideration:

- Sex
- Age – the age of the informants ranged from 15 to 65 and for the purpose of the statistical processing of the acquired data five categories were identified: <18, 19-24, 24-40, <40, <60.[1]
- Education – six categories were identified (the abbreviations refer to the graphs of the non-linear canonical analysis – see below): primary school (zs); some high school (gym), high school graduate, some college (ss); college graduate (bc); graduate school (vs).
- Occupation – the informants were grouped into four categories: 1 engineering, IT, health-care, scientific; 2 lawyers, journalists, teachers, administrative workers; 3 manual, artistic; 4 housewives, students, pensioners, unemployed.
- Language background – designing the categories within this factor proved rather complicated, which is why the typology cannot be presented within a few lines (as the factors above). Consequently, the following paragraphs describe the process of gaining, assessing and processing the data so that a typology could be established.

Each of the above-mentioned factors can raise many questions, yet that of the language background seems to present the most complex issue. The basic aim of the research was to compare the word-formation strategies in the Slovak and the English languages in those Slovak and English informants whose language background is Hungarian.

The analysis of the data showed a heterogeneous nature of the Hungarian-English group of informants. Most of them came from the families of Hungarian emigrants in English-speaking countries, in the majority of cases the USA or Australia. Two general tendencies were observed. First, the ancestors of the informants (or the informants themselves) mostly left their homes because of political persecution that was caused by their cultural background (e.g., they were of Jewish origin or Hungarians living in Romania). Consequently, their language background mostly included – in addition to English and Hungarian – also some other language. Secondly, the emigrants frequently found their life partners among other emigrants, very often of different origin, and in this way the language background of their children (our informants) consisted of English, Hungarian and some other language, e.g. Russian, Polish, Croatian, Rumanian, Spanish, Italian, etc.

On the other hand, the Hungarian-Slovak group of the informants was more homogeneous. They developed their bilingualism thanks to the historical background of the territory they came from – the majority of them had their roots in the southern part of Slovakia bordering on Hungary. This territory is well known for strong cultural and language bonds to the Hungarian language.

For the sake of statistical evaluation, Sapir's (1921) morphological typology was adopted. The reason for this was that the problems of word-formation typology and word-formation universals have been rather neglected in morphological/typological research. By implication, no word formation typology has been developed yet. In Sapir's typology, synthetic/inflective languages (e.g. Slovak) are characterised as languages in which grammatical relationships are expressed by inflection; synthetic/agglutinative languages (e.g. Hungarian) make use of agglutination, and analytic/isolating (e.g. English) express grammatical relations by word

---

[1]   The grouping of informants according to age was consulted with a distinguished Slovak sociolinguist Slavomir Ondrejovic. The age limits were determined by the age of informants – the youngest were about 16, the oldest 70. These limits were caused especially by 2 factors. Firstly, the Internet skills – since the questionnaire was distributed mainly by means of e-mail communication; secondly the cognitive abilities of informants.

order. Therefore, in view of our research objectives, the following language typology was used:

(6)   synthetic/inflective (SF) – e.g. Slovak, German, Russian, Croatian, Czech;
         synthetic/agglutinative (SAg) – e.g. Hungarian;
         analytic/isolative (AI) – e.g. English, French, Romanian, Italian, Spanish.

Based on this information, the following language groups were established:

(7)   SAg+SF+AI
         SAg+AI+SF
         SF+AI+SAg
         SF+SAg+AI
         AI+SAg+SF
         AI+SF+SAg
         AI+SAg
         SAg+AI
         SF+SAg
         SAg+SF.

The informants were grouped according to (a) their bilingualism, and (b) self-evaluation of their language skills. They were asked to evaluate

(8)
- their own language skills
- their parents' language skills
- the language used in their household

Letters A – E were used to mark the specific level, with A indicating fluency, and E rather poor level of language skills.

The order of the languages in (5) indicates the level of the language skills of the individual informants. For example, Hungarian-Slovak informants were integrated into the group SF (synthetic/inflective) + SAg (synthetic/agglutinative) if they indicated the information about the language background in the following way:

(9)   INFORMANT

Language	Level
Slovak	A
Hungarian	B

FATHER

Language	Level
Hungarian	A
Slovak	D

MOTHER

Language	Level
Slovak	A
Hungarian	B

LANGUAGES AT HOME

Language	Level
Slovak	A
Hungarian	B

Table 1 and Graph 1 provide the structure of informants for English questionnaires according to their language background, Table 2 and Graph 2 provide the same structure for Slovak questionnaires.

Language type	No. of informants
SAg+SF+A	4
SAg+AI+SF	27
SF+AI+SAg	3
SF+SAg+AI	2
AI+SAg+SF	8
AI+SF+SAg	3
AI+SAg	24
SAg+AI	75
Overall	146

Language background
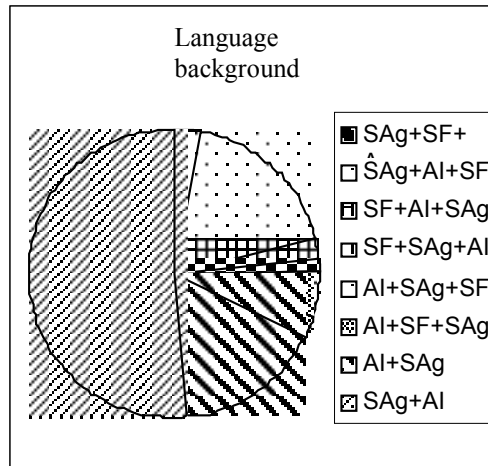
Table 1   Graph 1

Language type	No. of informants
SAg+SF+AI	37
SF+SAg	12
SAg+SF	93
Overall	142

Language background
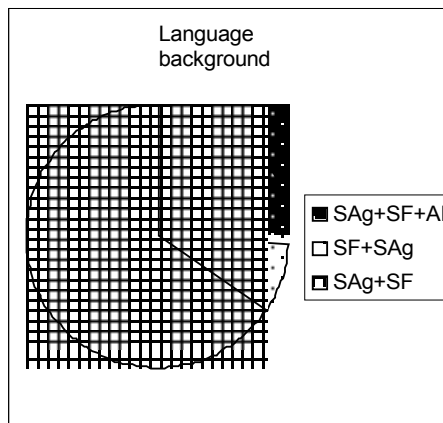
Table 2   Graph 2

Table 3 compares the data for English and Slovak questionnaires, and Table 4 gives the same data in percentages:

Type	JP 1	JP 2	JP 3	JP 4	JP 5	JP 6	JP 7	JP 8	JP 9	JP10	Total
AJ	4	27	3	2	8	3	24	75	0	0	146
SJ	37	0	0	0	0	0	0	0	12	93	142
Total	41	27	3	2	8	3	24	75	12	93	288

Table 3

Type	JP 1	JP 2	JP 3	JP 4	JP 5	JP 6	JP 7	JP 8	JP 9	JP10	Total
AJ	2,74%	19,49%	2,05%	1,37%	5,48%	2,05%	16,44%	51,37%	0,00%	0,00%	50,69%
SJ	26,06%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	8,45%	65,49%	49,31%
Total	14,24%	9,38%	1,04%	0,69%	2,78%	1,04%	8,33%	26,04%	4,17%	32,29%	100,00%

Table 4

Legend:[2]
JP – Language background
JP 1 –   SAg+SF+AI
JP 2 –   SAg+AI+SAg
JP 3 –   SF+AI+SAg
JP 4 –   SF+SAG+AI
JP 5 –   AI+SAg+SF
JP 6 –   AI+SF+SAg
JP 7 –   AI+SAg
JP 8 –   SAg+AI
JP 9 –   SF+SAg
JP 10 –   SAg+SF

It follows from Tables 1-3 and graphs 1-2 that the language background of English informants is more diverse than that of Slovak informants, which naturally follows from the country of their origin. If we add up all three possible combinations of three language types (AI+SAg+SF, AI+SF+SAg, SF+SAg+AI, ...) the number of English questionnaires is 48, while there are only 37 Slovak questionnaires falling within these combinations. A combination of two language types for English questionnaires (AI, Sag) occurs in 98 cases, and the corresponding combination of two language types for Slovak questionnaires occurs in 105 cases.

If English and Slovak informants are compared from the point of view of language background the Slovak sample is much more homogeneous. All Slovak informants adduce Slovak and Hungarian. The two languages are rarely completed with another language – in contrast to the situation in the English sample – mostly German, English, Russian and Czech occur as a third language. These are mostly languages taught at school, and thus not affecting the status of our informants as bilingual speakers. In spite of this fact, it is this sample of informants that frequently made use of English words and/or suffixes.

The questionnaires, designed in two languages (Slovak and English), consisted of two parts. The first part examined the naming strategies, the second part collected selected sociolinguistic data.

---

[2]   The abbreviations (e.g. AJ, SJ, JP) are based on the Slovak language for the reason that the data were statistically processed by a Slovak software.

## 4.2. Analysis of Word-Formation Strategies

The initial part of the questionnaire consisted of various tasks with one basic aim – the informants were supposed to coin new, potential complex words denoting Agents. The first task was based on motivation by words. The informants were provided with several possibilities of how to name a person, an Agent performing an action, e.g.:

(10)   A person who produces yogurts:

a)	yogurter	i)	yogurtie
b)	yougurtor	j)	yogurtman
c)	yogurtent	k)	yogurt-producer
d)	yougurtier	l)	yogurt-person
e)	yougurtist	m)	yogie
f)	yougurtitor	n)	yoducer
g)	yogurtnik	o)	other
h)	yogurster		

The informants' task was to select one of them, in their view the most appropriate name for such a person.

The second task made use of visual motivation. The informants were asked to name the Agent in the picture, for example:
How would you name this person?

(11)
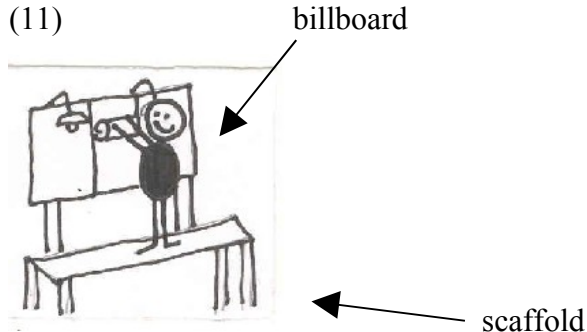
The third task consisted of a description of a non-existing game and of its playground layout. Based on the given description the informants were asked to name the players involved in the game:

(12)   In the middle of the playground, there is a basket with tennis balls. The balls are in three colours and each ball has its value. Among them, there is a golden ball with the highest value. The playground is divided into two halves – one for each team. There is a basket at both ends of the playground. The baskets look like basketball baskets but they have a bottom. The aim of the game is to shoot the balls in the basket placed in the middle of the playground into the baskets placed at the end of the opponent's playground. The points are counted according to the colour of the shot balls. The game finishes in the moment when all balls from the basket in the middle are shot or when one team succeeds in shooting the golden ball in the opponent's basket.
Each team has six players. Player 1 takes the balls out from the basket in the middle of the playground. Players 2 and 3 have tennis rackets and their task is to strike the ball passed by player 1 into the opponent's basket. Players 4 and 5 defend with tennis

rackets the team's basket at the end of the playground. Player 6 picks up the balls passed by players 4 and 5 during the defence as well as the balls that get to his part of the playground from the opponent's playground and bats them to players 2 and 3.
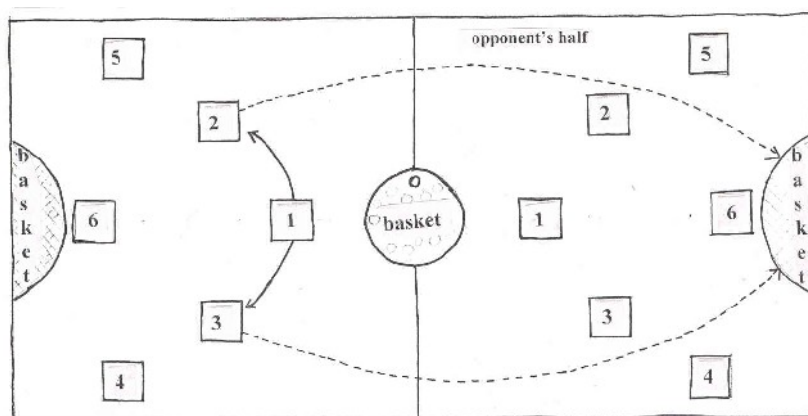
The task consisted in giving names to the individual players

Not all questionnaires were filled out completely. This is illustrated in Table 5

	English questionnaires	Slovak questionnaires
Total Number of questionnaires	170	158
Total number of questionnaires analysed	146	142
Ideal number of complex words	1606	1562
Actual Number of complex words	1252	1195

Table 5

Complex words selected/proposed by the informants were analysed from the viewpoint of onomasiological types, morphological types and word-formation types. Their productivity was calculated, and correlations between the sociolinguistic factors and productivity were searched for. The primary aim was to find out the level of the influence of the sociolinguistic factors on productivity as reflected in the preferred naming strategies. The total productivity was compared to the productivity at individual levels in relation to each of the sociolinguistic factors. Furthermore, the two groups of bilingual informants were compared. The statistical programmes Statistica and SPSS, including non-linear canonical correlation, were applied.

5. Research Results

Various statistical methods (Statistica SPSS, canonical correlation, non-canonical correlation) were used in our research.

5.1. Results of the Non-Linear Canonical Analysis

In our research, two correlation methods were employed. In general, canonical correlation is used to study the relation between two sets of variables (e.g. age and language background). On the other hand, the non-canonical correlation enables to search for relations between more than two sets of variables. An important advantage of the non-linear canonical correlation is that individual variables can be nominal, ordinal and interval, as opposed to canonical correlation, which makes it impossible to work with more than one set of variables.

When evaluating the data, we searched for the strongest correlation among the individual sociolinguistic factors. The results showed that the sociolinguistic factors of education and occupation bore the required correlation; in further analysis, they were approached as one variable. At the same time, the factor of the language background proved to be fully independent of sex, age, education and occupation. At the beginning of our analysis, Excel tables were used. Since the non-linear canonical correlation works with a scope from one, it turned out that Excel tables were not suitable for the non-linear canonical correlation due to strong presence of zeroes. For this reason, each data was considered separately.

The results are provided in the following graphs, where the Slovak word *pohlavie* stands for sex; *vek* for age; the abbreviated form *vzdel* means occupation; *JP* language background. JP1 – JP8 are the language background categories that were identified based on the informants' self-evaluations:

(13) JP1=SAg+SF+AI
JP2=SAg+AI+Sag
JP3=SF+AI+SAg
JP4=SF+SAg+AI
JP5=AI+SAg+SF
JP6=AI+SF+SAg
JP7=AI+SAg
JP8=SAg+AI
JP9=SF+SAg
JP10=SAg+SF

The abbreviations OT, WFT, and MT stand for the onomasiological type, word formation type, and morphological type, respectively. The characteristics of the onomasiological types (from OT1 to OT5) were briefly introduced in Chapter 3. Moreover, the original lists of onomasiological types and word-formation types were completed with OT6 and WFT6, representing borrowings. An overview of the WFTs and MTs is as follows:

(14) Word formation types:
WFT1: Object – /Action/ – Agent   WFT4: Object – Instrument – Agent
WFT2: Object – Action – Agent   WFT5: Others
WFT3: Action – Agent   WFT6: Result – /Action/ – Agent

(15) Morphological types:
MT 1: S + suffix   MT 3: S + S + suffix   MT 5: Conversion

MT 2: S + S   MT 4: Others

The following comments mainly pay attention to the parameter of language background (JP), in particular, to JP1 because it was present in both groups of informants. It represents the language background with a stronger Hungarian language; one synthetic/inflective and one analytic/isolative language.

5.2. Non-Linear Canonical Correlation of the Onomasiological Types and Sociolinguistic Factors
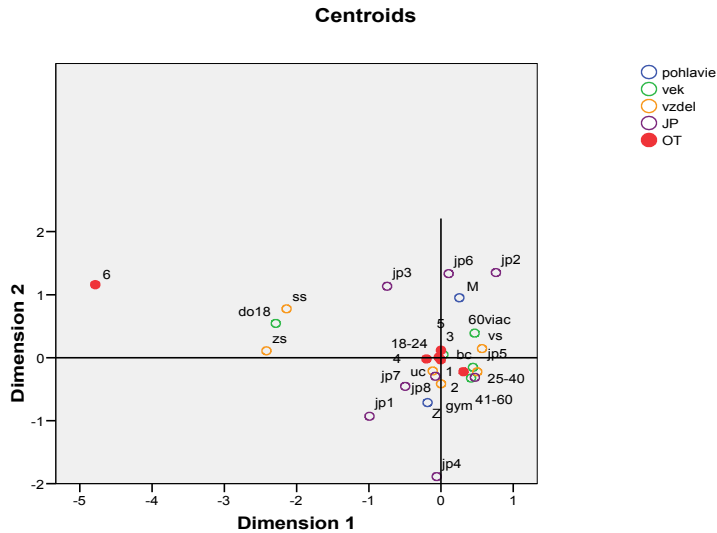
Graph 3 *English questionnaires*
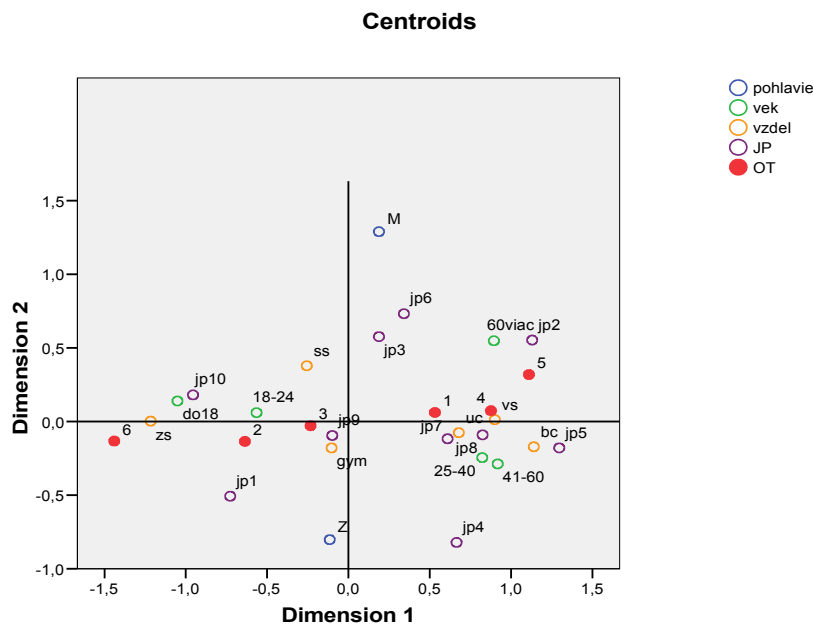
Graph 4 *Slovak questionnaires*

Centroids are intersections of Dimensions 1 and 2. These represent specific sociolinguistic factors and the productivity of onomasiological types. The graphs allow us to search for various correlations and the amount of possible information given in the graphs is considerable. For illustration, let us compare the influence of Hungarian language in both groups of informants (English-Hungarian and Slovak-Hungarian) on the productivity of onomasiological types and word formation types (in the graphs, the productivity of OTs, WFTs and MTs are marked with red dots and numbers).

JP1 represents the language background with strongest Hungarian background (SAg+SF+AI). Since the red onomasiological type dots in the English graph are concentrated in a cluster, which causes some problems with the data interpretation, it will be more advantageous to start with Slovak graphs. For JP1 in graph 4 (Slovak questionnaires), the closest onomasiological type is Onomasiological Type 2, which means that the correlation between JP1 and Onomasiological Type 2 is the strongest of all. Similarly, for JP1 in graph 3 (English questionnaires), the closest onomasiological type is Onomasiological Type 1, closely followed by Onomasiological Type 2 and Onomasiological Type 4. This comparison enables us to assume that Hungarian as a background language increases the productivity of Onomasiological Type 2. Furthermore, in the English graph, Onomasiological Type 1 is closer to JP1. It means that the correlation between these two variables is stronger than that between JP1 and OT2/OT4. It can be caused by heterogeneous language background of the English-Hungarian group of informants. However, the same influence of the Hungarian language on Onomasiological Type 2 in both groups of informants is undisputable.

As a next step, let us compare the most frequent language background types in both language groups of informants – the language backgrounds JP8 (Hungarian + English) and JP10 (Hungarian + Slovak). In both language backgrounds, the Hungarian language is the stronger one. The Slovak informants with this background preferred borrowings and Onomasiological Type 2. The English informants made use of Onomasiological Type 1 (the most productive one in English complex words) or Onomasiological Type 4.

The correlations between the onomasiological types and sociolinguistic factors in Slovak complex words confirmed the previous results – borrowings were preferred by pupils and students under 18 years of age, with Hungarian as the stronger background language background, and Slovak as a weaker language. Informants aged 18-24, with a secondary grammar school education and the language background of SF + SAg (stronger Slovak, weaker Hungarian), used the most productive onomasiological type (OT3) in the Slovak language.

The graphs also offer the possibility to compare the influence of the individual language backgrounds on the productivity of onomasiological types, word-formation types, and morphological types. A good example is the language background JP1, since it was present in both groups of informants. It represents a combination of the Hungarian language, which is the strongest, and a synthetic/inflective language and an analytical/isolative language. In the Slovak complex words, it mainly influenced the productivity of Onomasiological Type 2. A similar tendency could be observed in the English complex words, although this language background also correlated with Onomasiological Type 4. It is assumed that Hungarian language influences the productivity of Onomasiological Type 2.

5.3. Non-Linear Canonical Correlation of the Word-Formation Types and Sociolinguistic Factors
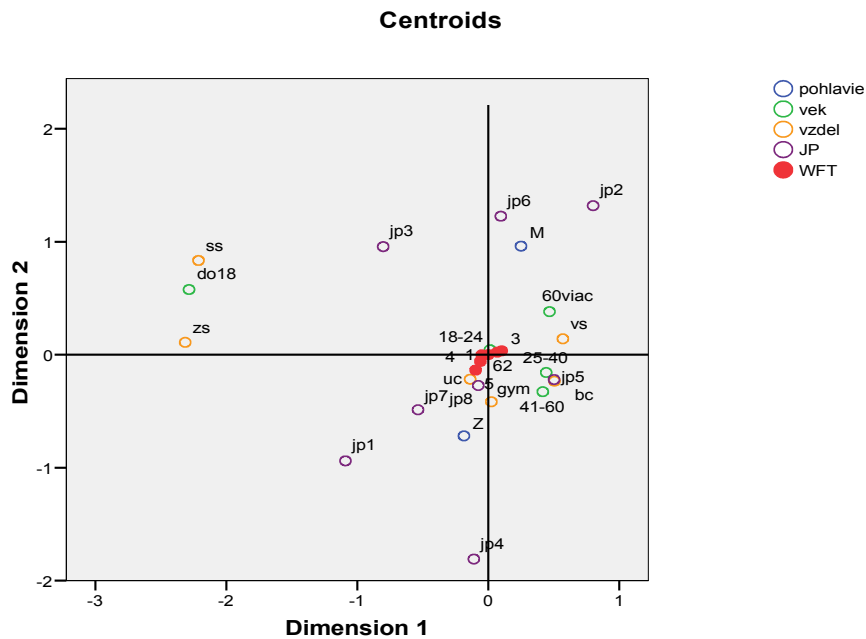
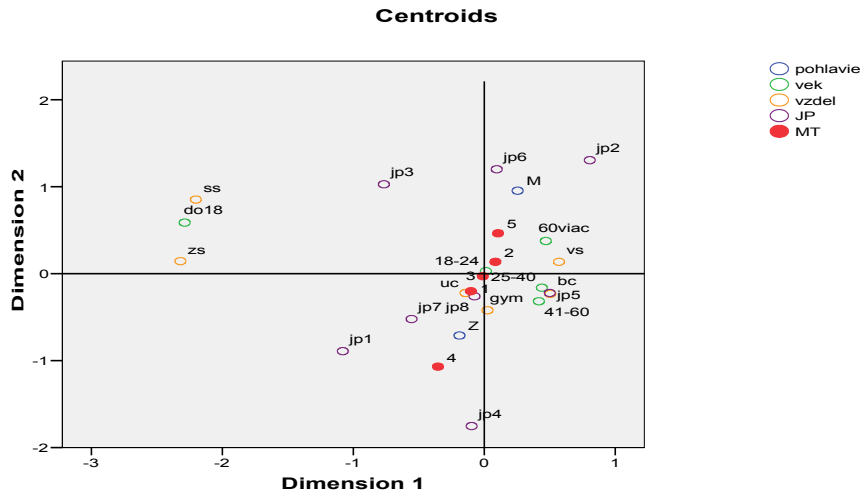Graph 5 *English questionnaires*
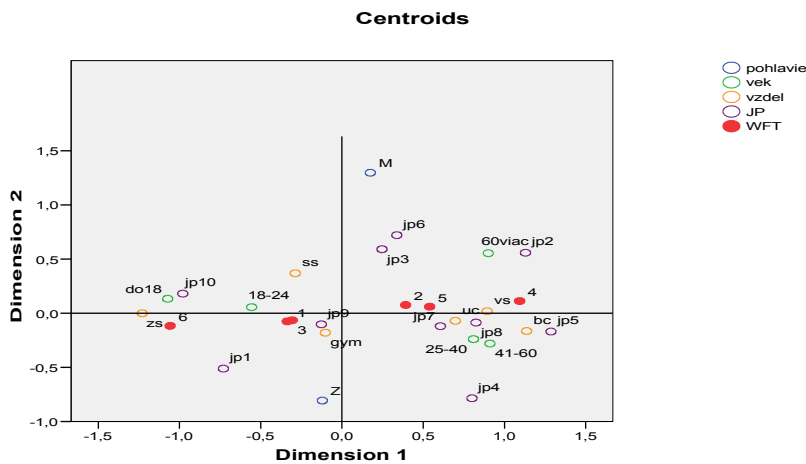
Graph 6 *Slovak questionnaires*

For JP1 in graph 6, the closest word-formation types are WFT1 and WFT3. It means that JP1 exerts pressure on WFT 1 and WFT 3. Even though the English graph is difficult to interpret due to the WFT cluster, it is obvious that WFT 1 is the closest of all word-formation types. By implication, Hungarian language increases the productivity of the same WFT in both groups of informants.

Graphs 5 and 6 show the correlation between the word-formation type and sociolinguistic factors in English and Slovak. In graph 5, the pressure of JP1 on Word-Formation Type 5 (the most productive word-formation type in English complex words) is visible. On the other hand, while in graph 4, JP1 is quite far from the central axis, its pressure on Word-Formation Type 1 and Word-Formation Type 3 (the most productive word-formation types in the Slovak

language) is noticeable.

5.4. Non-Linear Canonical Correlation of the Morphological Types and Sociolinguistic Factors

Graph 7 *English questionnaires*
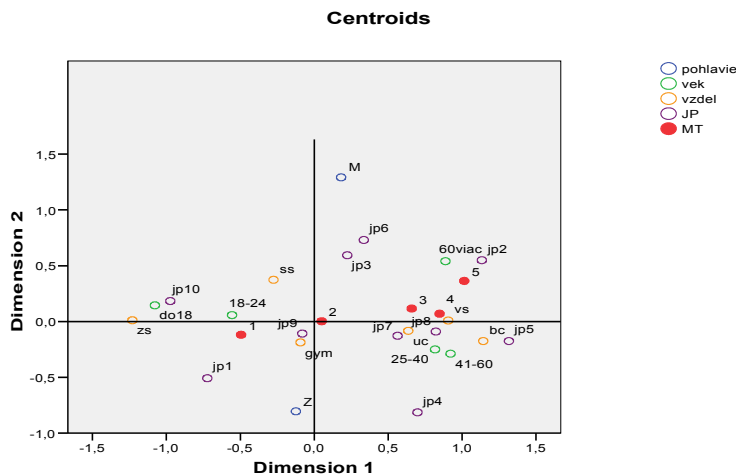
Graph 8 *Slovak questionnaires*

In the English complex words, the most productive morphological type was "Stem + Stem + Suffix". It was preferably used by informants aged 18-24, of high school education, and with language background combining the Hungarian and the English languages. In the Slovak complex words, the most productive morphological type was "Stem + Suffix" that was used by those of the same age and education. As for the Slovak language the only difference concerns the language background – the change in the language background causes the change in the productivity of the morphological type.

5.5. The Sociolinguistic Factors, New Complex Words and Productivity

Based on the results, it is possible to arrive at the following conclusions:

5.5.1. Sex

Neither English nor Slovak data showed significant influence of sex on the choice of the onomasiological type. Since this result was confirmed for both groups of informants it can be assumed that the influence of sex on productivity in word formation is not relevant, especially in comparison with the factors of age, education and occupation. The analysis of the influence of sex on word-formation types and morphological types brought the same result.

5.5.2. Age

The influence of age was unequivocal. The lower the age of the Hungarian-English informants the stronger the tendency towards complex words with simple onomasiological structure or non-transparent complex words. At the same time, the Slovak-Hungarian informants of a younger age preferred borrowings that are too non-transparent in the Slovak language. The influence of age was also observed at the level of word-formation type – the younger age categories prefer word-formation types that were not typical of the given conceptual category. It can be explained as their effort at originality. This phenomenon was observed in both groups of informants and it can be generalised as a phenomenon typical of the relationship between the word-formation type and age.

To sum up, the sociolinguistic factor of age influences the naming strategies in the process of coining new complex words. The most striking deviation from the norm is observable in the age category 18 – 24. These informants differ from other age categories especially in the preferred onomasiological type, word-formation type and morphological type. The types chosen by them are not very productive in other age categories.

5.5.3. Education

The non-linear canonical correlation showed strong association between the factors of age and education. By implication, the analysis of the relationship between the factors of education and productivity in word-formation displayed similar results. The influence of education was the most visible at the level of onomasiological types. The higher the education of the informants the stronger tendency towards a more transparent onomasiological structure of the coined complex words.

5.5.4. Occupation

The influence of occupation was in accordance with the influence of age and education, since the factor of occupation highly correlated with them. From the perspective of occupation, the most creative group in the field of word-formation seems to be the group of students. Both Slovak and English students' word-formation strategies tend to deviate from expected ones. Similar deviations can be observed in the category of manual workers and artists.

5.5.5. The Influence of the Language Background

The non-linear canonical correlation confirmed the fact that the language background is a completely independent sociolinguistic factor different from the rest of the sociolinguistic factors. While strong correlations were found between the factors of age and education, education and occupation, and occupation and age, no similar correlations were identified for the language background.

All statistical methods clearly confirmed the influence of the language background on naming strategies. While the productivity of onomasiological types for Slovak and English differed onomasiological types 1 and 3 appeared among the most productive onomasiological types in both languages. It is beyond dispute that the Hungarian language shared by both groups of our respondents may be held responsible for this similarity. The canonical correlation confirmed this observation, too. In addition, in both groups of informants the following tendency was observed: the stronger the influence of the Hungarian language the higher the productivity of Onomasiological Type 2, which indicates that the Hungarian language does not favour a more detailed motivation in the naming strategies; instead it appears to prefer the brevity of expression.

The pressure of the Hungarian language was more apparent at the level of word-formation types – the productivity results in both groups were nearly identical. The non-linear canonical correlation revealed the role of language background also for the level of morphological types. In summary, the influence of the language background is obvious especially at the level of the onomasiological and word-formation types. A low value of 'p' in the canonical correlation proves the statistical significance of the research. All in all, our results confirm the hypothesis of the influence of the language background on productivity in word formation.

6. Conclusions

The research results confirmed the hypothesis of the influence of sociolinguistics factors upon the naming strategies. No doubt, new complex words come into existence at the crossroads of the sociolinguistic factors and the pressure of productive onomasiological types, word formation types, and morphological types. This pressure was the most visible at the level of the onomasiological and word-formation types. The sociolinguistic factor of the language background turned out to be an independent and autonomous sociolinguistic factor.

The strongest influence of the sociolinguistic factors was observed for age and occupation. The least significant influence was identified for the factor of sex. The influence was the best observable at the level of onomasiological types and word-formation types. The strongest tendency was the correlation of students (aged 18-24) with some high school education and with the language background SF + AI (stronger Slovak, weaker Hungarian). This group of Hungarian-Slovak informants, instead of coining new complex words, preferred to use borrowings in the Slovak language.

The research suggests that the influence of the sociolinguistic factors is significant especially at the level of onomasiological types and word-formation types. The influence was less visible at the level of the morphological types. According to the canonical correlation the level of word-formation types seems to show great potential for further investigation in the field of word formation. In addition, a more homogeneous Slovak group of informants showed strong correlation between the language background, on the one hand, and onomasiological type and word-formation type, on the other.

Lívia Körtvélyessy
Department of British and American Studies
Faculty of Arts
P. J. Šafárik University
Petzvalova 4
04011 Košice, Slovakia
livia.kortvelyessy@upjs.sk

References

Baayen, Harald (1989), *A Corpus-Based Approach to Morphological Productivity: Statistical Analysis and Psycholinguistic Interpretation*, doctoral dissertation Free University of Amsterdam.

Baayen, Harald (1992), *A Quantitative Approach to Morphological Productivity*, in Booij, Geert / van Marle, Jaap (eds.), *Yearbook of Morphology 1991*, 109-149, Dordrecht: Kluwer.

Baayen, Harald (1993), *On Frequency, Transparency and Productivity*. In Booij, Geert / van Marle, Jaap (eds.), *Yearbook of Morphology 1991*, 181-208, Dordrecht: Kluwer.

Baayen, Harald R. / Lieber, Rochelle (1991), "Productivity and English Derivation: A Corpus-Based Study", *Linguistics* 29: 801-804.

Bauer, Laurie (1983), *English Word-Formation*, Cambridge: Cambridge Univ. Press.

Bauer, Laurie (2001), *Morphological Productivity*, Cambridge: Cambridge Univ. Press.

Černý, Jiří (1996), *Dějiny lingvistiky*, Olomouc: Votobia.

Dokulil, Miloš (1962), *Tvoření slov v češtine,* Praha: Nakladatelství československé akademie.

Downes, William (1998), *Language and Society*, Cambridge: Cambridge Univ. Press.

Fabb, Nigel (1988), "English Suffixation Is Constrained Only by Selectional Restrictions", *Natural Language and Linguistic Theory* 6: 527-539.

Gaeta, Livio (2005), "Word Formation and Typology", in: Booij, Geert et al. (eds.), *On-line Proceedings of the Fourth Mediterranean Morphology Meeting, Catania*. [online]. 157-169. Retrieved from: **http://morbo.lingue.unibo.it/mmm/mmm-proc/MMM4/157-169-Gaeta-MMM4.pdf** [access date: 2006-07-04]

Horecký, Ján (1983), *Vývin a teória jazyka.* Bratislava: SPN.

Horecký, Ján (2000), "Človek a jeho jazyk: Esej čiže pokus o charakteristiku vzťahu", *Jazykovedný časopis* 51: 3 -6.

Kastovsky, Dieter (1986), "The Problem of Productivity in Word-Formation", *Linguistics* 24: 585-600.

Marle, Jaap van (1986), "The Domain Hypothesis: The Study of Rival Morphological Processes", *Linguistics* 24: 601-627.

Plag, Ingo (1999), *Morphological Productivity.* Berlin: Mouton de Gruyter.

Rainer, Franz (1993), *Spanische Wortbildungslehre*. Tübingen: Niemeyer.

Rainer, Franz (2005), *Constraints on productivity*, Stekauer, Pavol / Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 335-52, Dordrecht: Springer.

Sapir, Edward (1921), *Language,* New York: Harcourt, Brace and World.

Štekauer, Pavol (1998), *An Onomasiological Theory of English Word-Formation*, Amsterdam/Philadelphia: John Benjamins.

Štekauer, Pavol (2005), "Onomasiological Approach to Word-formation", Štekauer, Pvaol / Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 207-232, Dordrecht: Springer.

Štekauer, Pavol et al. (2005), "Word-Formation As Creativity Within Productivity Constraints: Sociolinguistic Evidence", *Onomasiology Online* 6: 1-55.

Appendix 1

Questionnaire

Dear respondent,

our aim is to find out the way of forming new words in English language. To reach our aim we have prepared the following questionnaire and we would like to ask you to spend few minutes on filling it out.

The questionnaire is based on words that do not exist in English language. It is a questionnaire, not a test – there are no "correct" answers. We simply want to find out what words you think will fit best the described situations.

Task 1. Choose the word that you think is the most suitable for the person described in the question.

1. A person who produces yogurts:

a) yogurter

b) yogurtor

c) yogurtent

d) yogurtier

e) yogurrtist

f) yogurtitor

g) yogurtnik

h) yogurtster

i) yogurtie

j) yogurtman

k) yogurt-producer

l) yogurt-person

m) yogie

n) yoducer

o) other

2. A person attending the scales in the baggage room.

a) scaler

b) scaleman

c) scaleperson

d) scalist

e) scalie

f) scalnik

g) baggageman

h) baggage-scaler

i) baggage-person

j) baggager

k) other

3. A person who is engaged in the research of comets.

a) Cometer

b) Cometist

c) Cometant

d) Cometitor

e) Cometnik

f) Cometie

g) Cometman

h) Comet-researcher

i) Comsearcher

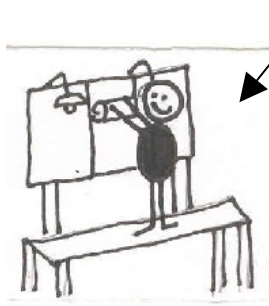j) Comet-explorer

k) other

4. A person interested in Harry Potter:

a) Potterfan

b) Potterist

c) Potterperson

d) Potterphil

e) Potterman

f) Potterant

g) other

5. A person dependent on phoning:

a) Phoner

b) Phonnik

c) Phonist

d) Phonant

e) Phonee

f) Phone-obsessee

g) Phone-obsessive

h) phoneman

Task 2. Name people in the following pictures

1.   billboard

scaffold

How would you name this person?

2.

newspapers

How would you name this person ?

Task 3. Read the following description of a sport game.

In the middle of playground there is a basket with tennis balls. The balls are in three colors and each ball has its value. Among them there is a golden ball with highest value. The playground is divided into two halves - one for each team. There is a basket at both ends of the playground. The baskets look like basketball baskets but they have a bottom. The aim of the game is to shoot the balls in the basket placed in the middle of the playground into the baskets placed at the end of the opponent's playground. According to the color of the shot balls the points are counted. The game finishes in the moment when all balls from the basket in the middle are shot or when one team succeeds in shooting the golden ball in the opponent's basket.
Each team has six players.
Player 1 takes the balls out from the basket in the middle of the playground.
Players 2 and 3 have tennis rackets and their task is to strike the ball passed by the player 1 in to the opponent's basket.

Players 4 and 5 defend with tennis rackets the team's basket at the end of the playground.
Player 6 picks up the balls passed by players 4 and 5 during the defense as well as the balls which get to his part of the playground from the opponent's playground and bats them to players 2 and 3.
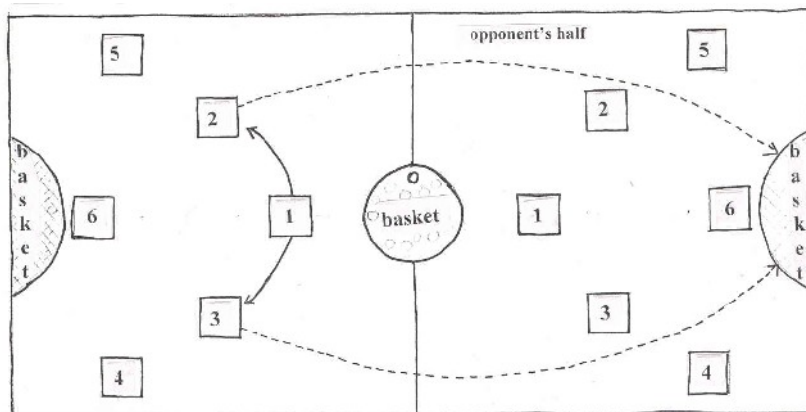
Name the player 1.

Name the players 2 and 3.

Name the players 4 and 5.

Name the player 6.

- - - - - -

SOCIOLINGUISTIC DATA

These data will be used for research only, it will not be used to identify any individual. The data are important for our research. Fill your data in and return the questionnaire, please.

A. PERSONAL DATA

Age: Sex: Occupation:

Where born (state or country):

Your education (circle the highest level that applies)

a) *Some high school*   d) *College Graduate*

b) *High School Graduate*   e) *Graduate School*

c) *Some College*

B. LANGUAGE BACKGROUND

Please read the following scale describing various levels of language knowledge.

A. I am a native speaker.

B. I am not a native speaker, but native speakers usually think I am.

C. I speak the language fluently, but I have an accent or sometimes say things that do not sound natural to native speakers.

D. I can speak the language fairly well, but sometimes I have to hesitate to think of words or grammatical constructions.

E. I know a little bit, but I have a hard time conversing normally in the language.

Think about the languages you speak and match them with the above mentioned levels A-E.

Language   Ability (A, B, C, D, E)

1.

2.

3.

C. YOUR FATHER'S LANGUAGE BACKGROUND

Please list the languages that your father speaks and rate his ability to the following scale.

A. He is a native speaker.

B. He is not a native speaker, but native speakers usually think that he is. (Near-native)

C. He speaks the language fluently, but he has an accent or sometimes says things that do not sound natural to native speakers.

D. He can speak the language fairly well, but sometimes he has to hesitate to think of words or grammatical constructions.

E. He knows a little bit, but he has a hard time conversing normally in the language.

Language   <u>Ability</u> (A, B, C, D, or E)

1.

2.

3.

D. YOUR MOTHER'S LANGUAGE BACKGROUND

Please list the languages that your mother speaks and rate her ability according to the following scale:

A. She is a native speaker.

B. She is not a native speaker, but native speakers usually think that she is. (Near-native)

C. She speaks the language fluently, but she has an accent or sometimes says things that do not sound natural to native speakers.

D. She can speak the language fairly well, but sometimes she has to hesitate to think of words or grammatical constructions.

E. She knows a little bit, but she has a hard time conversing normally in the language.

<u>Language</u>   <u>Ability</u> (A, B, C, D, or E)

1.

2.

3.

E. LANGUAGES SPOKEN IN YOUR HOME

1._____

      a) always   c) occasionally

      b) frequently (daily or nearly so)   d) almost never

                                                      e) never

2._____

      a) always   c) occasionally

      b) frequently (daily or nearly so)   d) almost never

                                                      e) never

3._____

      a) always   d) almost never

      b) frequently (daily or nearly so)   e) never

      c) occasionally

May we contact you for help in locating other people who might be willing to complete this survey?
Y/N

Your contact information (name, address, phone, e-mail):