**ЦЕННОСТЬ ВЛАДЕНИЯ ВТОРЫМ ЯЗЫКОМ
НА РЫНКЕ ТРУДА**

Н.Н. Хиоарэ

*В настоящее время на рынке труда заработная плата специалистов зави-сит
от уровня их компетентности в определенной области. Основным состав-*

ляющим компонентом данного уровня является не только знание базовой лекси-ки (повседневного языка), но и знание профессиональной лексики (в рамках дан-ной статьи речь идет о терминах в области экспорта и импорта). Автор предлагает модели изучения терминов близкородственных и дальне-родственных языков.

Ключевые слова: родственные пары, дальние языковые пары, связующий язык, модель внутриродственного перевода, модель межродственного перево-да, наследственная корневая форма.

THE VALUE OF SECOND LANGUAGE SKILLS
ON LABOUR MARKET

N.N. Hioară

The modern labor market pays employees according to their level of competence in a certain area. A component of this level is knowledge of several languages, moreo-ver, knowledge of not only the main lexical fund (everyday language) but also the terms of a specific science (here the terms of export and import). The author proposes a model for studying the terms of closely related and distantly related languages.

Keywords: cognate pairs, distance pairs, bridge language, intra-family transla-tion model, cross-family translation model, ancestral root form.

When the experts evaluate the costs of investment in second languages as ele-ments of human capital there arises a number of difficulties. Individuals getting an education usually endure some costs: those connected with expenditure on books, tuition, video films, etc., and, naturally forgone expenditures (go without something desirable). The first type of cost involves direct expenditure on school materials, though not nil, being relatively minor and can be assumed away (that meanwhile would not result in a major difference in the estimated rates of return in the nearest or in the remotest future). The second cost that concerns forgone expenditures, in terms of the cost for second language skills, tend to be zero for learners under legal working age.

Moreover, the latter expenditures would not be allowed to sell time for higher earnings, not spent at school. Consequently,the forgone expenditures can't be traded for a wage on the labour market, because there was no time specifically taken away, from language classes. The same is adequate for those who paid for evening classes because continuing education was arranged, and was done privately, it was not a state-imposed policy.

There is no data on the component of education spending provided by the state and therefore the current education accounting practices fall short in terms of analyti-cal data (the aggregate spending figures one may derive from Switzerland's language teaching cost. Approximately 10% of total education spending is devoted to second-

language, teaching. According to our estimation, a 'short stream' education covers 4,5% and a 'long-stream' education – 7%). The cost of language education should not be different in various linguistic environments. Just on the contrary the educational policy should embrace limitless diversity, although advantages and drawbacks (benefits or costs) are not confined to monetary return.

The linguistic policy should be considered an ill-advised economic standpoint since it underestimates the benefits and overestimates the cost of diversity. Both the benefits and the costs vary when the picture covers different linguistic environment.

Now turning from expenditures on linguistic skills, we try to direct to the aggregate data dealing with "earnings" or the so-called "private rates of return" to second language (SL) skills. The survey data including information about 1) earnings, second-level skills, education, and age are not collected by the national census (the Canada census being notable an exception) conform to the supposition that they are expensive and technically demanding. The data includes four skills (understanding, speaking, reading, and writing) and for each skill (separately) four skill levels (fluent, good, basic, none), non-school of L2 acquisitions, L2 use in the workplace, and standard social characteristics covering labour income. The data of the entire survey calculate the average earnings of groups in terms of their competence, and choices (active, receptive, oral, or written competence). The result reveals that there is a direct correspondence (a strong correlation) between earnings and competence.

This survey was provided by the potential employees using not reported income but the full-relevant equivalent of reported income.

Better paying jobs, higher education, better-educated people (some data include whether the applicants have had more Latin at school than others that don't necessarily conclude that they are rewarded for some skills to translate Renaissance authors), specialized courses in certain fields of economics are priorities to be taken on in order to obtain high earnings.

We tried to go on with the abilities dealing with language skill education and started to work with economic terminology for export and import (to and from Russia, Ukraine, Romania, Belarus, and Germany).

The scientists' efforts are directed at working out certain methods (models) that will accelerate the assimilation of the most difficult lexical (or grammatical) element, by using the principle of conscious learning aimed at training specialists (experts) on the regional and global levels. The proposed approach (model) represents a method for inducing translation lexicons (terminology) based on transduction models of cognate pairs via bridge languages (in our paper the bridge language is English). Bilingual lexicons within language families (see Germanic Languages in the North: Icelandic, Norwegian, Faeroese, Swedish, Danish; in the West: English, Frisian; in the South: Flemish, Afrikaans, Dutch, Yiddish, German, and Gothic in the East) are induced using probabilistic string edit distance models. Translation lexicons for arbitrary distant language pairs are then generated by a combination of intra-family translation models (as the Romance family of languages: Romanian, Italian, Sardinian, Portu-

guese, French, Spanish, Catalan, Occitan, Rhaetian) and one or more cross-family translation models (as, for example, a Germanic and Romance lexicon). The best translation, especially of terms, up to 95% exact match accuracy may be (is) achieved on the target vocabulary. The data of 30-68% was achieved on inter-family test pairs. Thus, substantial portions of translation lexicons (especially terms) can be generated (reached) accurately for languages even where no bilingual dictionary or parallel corpora may exist.

Although a translation lexicon is considered to be a mapping from words in the source language (SL) to words in the target language (TL), still for each word in the SL many (good) dictionaries provide one or more words in the TL, out of which not all the variants might be appropriate translations in some economic contexts (the "channel captain").

Many of the world's major language lexicons are available even online, but they might be often quite limited because of possible intellectual property constraints. To less-spread (or lower-density) languages the terms or the so-called translation lexicons usually exist as a hard-copy (on paper) dictionary, if at all.

While teaching terminology there usually exists a difficulty in creating a translation lexicon from scratch, because it requires time-consuming work by experts trained in both languages. Even partial dictionaries for lower-density languages would essentially decrease the translators' work solicited to erect machine translation systems.

The course of Terminology is aimed at mastering terms for AESM students, comparing the lexicons of arbitrary languages using models of cognate pairs, where words from two close languages (i. e. from the same intra-family languages) share both meaning and a similar surface form, arising when both words are derived from an ancestral root form (e.g. „*nepot*" in Romanian (Ro.), „**nephew**" in English (Eng.) or Ro. „*sora*", Ru. „*sestra*", Eng. „**sister**" (Buck, 1949), as well as Ro. „*părinţi*", Ru. «сестра», Eng. "**parents**", etc. The paper analyses the entire *Harmonized System of Export and Import Terms* for Trade provided by the RM with 5 countries: Russia, Belarus, Ukraine, Germany, and Romania. The system covers 18 categories of goods each with more than 10 subcategories.

Among terms, there are a lot of cognate pairs (common root) categorized in Section V – "*Mineral Products*" covered by Annex 23. External trade with Russia: Romanian-English cognate pairs: 1) Ro. Gaz de sonde (petrol) şi alte hidrocarburi gazoase. Engl. Petroleum gases and other gaseous hydrocarbons. 2) Ro. Uleiuri din petrol sau obţinute din minerale bitumate. Engl. Petroleum oils and oils obtained from bituminous minerals. 3) Ro. Ghips; anhidrit Engl. Gypsum; anhydrite.

In the first two sentences (or subcategories) the cognate pairs match more than 95% of accuracy, although the average match of inter-family test pairs is considered 30-68% of accuracy. In the second comparison of the two sentences the cognate pairs' accuracy power match is also more than 95%.

The discrepancy is constituted only by the conjunctions *"şi"* and "as well as the prepositions *"din", "from"*, respectively.

The third cognate pair has 100% accuracy, although the SL and the TL are of two different families of languages, the former is a Romance language and the latter – a Germanic one, i.e. this cognate pair constitutes an arbitrary distant language pair generated by a combination of inter-family translation models used for cross-family (vocabulary) dictionary.

Another harmonized category of terms included in Section V "*Products of chemical industries*" shows that out of seven items (articles), six of them share the same accuracy as in the previous Section V. Only a single (No7) item doesn't match the same accuracy because of cross-family divergence: E.g. 1. Ro. Medicamente , Engl. Medicaments; 2. Ro. Îngrăşăminte (fertilizare), Engl. Fertilizers; 3. Ro. Săpunuri, Engl. Soap; 4. Ro. Insecticide, rodenticide, fungicide, erbicide, Engl. Insecticides, rodenticides, fungicides, herbicides ; 5. Ro. Produse cosmetice şi de parfumerie ; Engl. Cosmetic and perfumery products ; 6. Ro. Carbonaţi; percarbonaţi; carbonat de amoniu tehnic; Engl. Carbonates; peroxocarbonates; commercial amonium carbonate; 7. Ro. Preparate antigel şi lichide pentru degivrare; Engl. Anti-freezing preparations and prepared de-icing fluids.

The lack of direct cognate pairs in the 7th example, namely, Ro. "*antigel*" and Engl. "anti-freezing" or Ro. "*degivrare*" and Engl. "**de-icing**" represents a cross-family difficulty in translation (lexicon) for the students who don't know other Indo-European languages.

Obviously, not all translations are easy to translate. Here interviews "*false friends*" the so-called "*deceptive cognates*" "treacherous cognates", (e.g. Ro. "genial" is translated into Engl. as "**blind**" (in "genial weather"), "*drăguţ*" (in "*a genial kid*"), "*jucăuş*" (in "**a genial dog**"), "*prielnic*" (in "**a genial atmosphere**", etc.), because historically related, they are distant enough to be challenging to the model (e.g. Engl. "**trivial button**" and Ro. "*floare de butonieră*").

Depending on how closely two languages are related, they may share more or fewer cognate pairs. Analyzing another section of the article designated for the external trade with Russia by the main "*Harmonized System*" (HS) categories of goods (Section1) one can identify a number of cognate pairs as well as cognate distance. E.g. 1. Ro. *Carne de bovine*/ Engl. Meat of bovine animals 2. Ro. *Carne de ovine*/ Engl. Meat of sheep 3. Ro. *Lapte şi produse lactate*/ Engl. Milk and dairy produce. The translation demonstrates that languages are often close within a common language family (both Romance and Germanic languages are of the Indo-European Family of Languages) so that cognate pairs between the two groups of translations are common or identical, and significant portions of the translation lexicon can be induced with high accuracy. In Section 1 only one cognate pair was depicted – Ro. „*bovine*" and Engl. „**bovine**".

There are several cognate pairs in Section 2 „*Vegetable Products*". E.g.1. Ro. *Mere, pere şi gutui, proaspete*, Engl. Apples, pears and quince, fresh 2. Ro. *Legume în stare proaspătă sau refrigerate*, Engl. Vegetables, fresh or chilled 3. Ro. *Caise, cireşe, piersici, prune, proaspete*, Engl. Apricots, cherries, peaches, plumb, fresh; 4. Ro.

Struguri, proaspeți sau uscați (stafide), Engl. Grapes, fresh or dried; 5. Ro. <u>*Semințe de floarea-soarelui,*</u> Engl. <u>Sunflower seeds</u>; 6. Ro. *Făină de grâu sau maslin*, Engl. Wheat and maslin flour

The cognate pair in the first translation lexicon (TL) or item is represented by Ro. pere/ Engl. Pears. In the second – there is no one; in the third – there are two cognate pairs: Ro. *cireşe* / Engl. **cherries**, and Ro. *prune* / Engl. **plumb**. In the forth TL lexicon there isn't a single cognate pair. The Ro. word „s**tafide**" is not lexicalized; there are two cognate pairs, in the fourth TL lexicon: Ro. *seminţe* /Engl. **seeds** and the next one is Ro. *floarea soarelui* / Engl. **sunflower**. The external trade terms included in section 2 enumerate five cognate pairs.

Section III entitled "*Animal or vegetable fats and oils*" represented by two translation lexicons as cognate pairs.

Ex. Ro. *Margarină, amestecuri alimentare de grăsimi,* Engl. <u>Margarine; edible mixtures or preparations of animal or vegetable fats;</u> Engl. Vegetable oils; Ro. *Uleiuri vegetale.*

Out of the first two items of translations four cognate pairs were registered: Ro. *margarină* / Engl. **margarine**, the second pair is Ro. *amestecuri* / Engl. **mixtures**, the next one, the third cognate pair is Ro. *vegetale*/ Engl. **vegetable,** and the fourth cognate pair is Ro. *ulei(uri)* / Engl. **oils**.

Here are a number of cognate pairs in in Romanian and English in section III: *băuturi/* **beverages**; *tutun* / <u>**tobacco**</u>; *vin*/**wine**; *preparate* **/** **preparations**; *fructe* / **fruit**, *alte* / **other**; *părţi* / **parts**; *plante* / **plants**; *alcool etilic nedenaturat /* **undenatured ethyl alcohol**; *lichioruri* / **liqueurs**; *băuturi spirtoase* / **spirituous beverages**; *sosuri* / **sauces**; *condimente* / **condiments**; *muştar* / **mustard**; *cereale* / **cereals**; *făinuri* / **flour**; *produse* / **products**; *patiserie* / **pastry (cook)**; *bere* / bear; *malţ* / malt; *ciocolată /* chocolate; *conţine cacao* / containing cocoa; *ţigări de foi* / *trabucuri, ţigarete /* **cigars, cigarillos, cigarettes**; *conserve* / **preserve**s; *zahăr* / **sugar**.

In section VIII „*Raw hides and skins, leather, fur skin*s" there were not identified cognate pairs. These words were historically distant enough and can't be challenging to model. These words don't share cognate pairs. Articles from section "**Wood**" IX such as **wood continuously shaped** (Ro. *lemn profilat*), **particle board and similar boar**d (Ro. *plăci aglomerate şi panouri similare*), **fibreboard of wood or other ligneous materials** (Ro. *panouri fibrolemnoase sau din alte materiale lemnoase*), **plywood, veneered panels and similar laminated wood** (Ro. *placaj, lemn furniruit şi lemn stratificat similar*) share only one cognate pair (similar / similar). Historically related they are distant enough and share the fewest number of cognate pairs.

The same "laboring car" occurs in section X: "*Pulp of wood; paper and paperboard*". Among the translation lexicons as English **uncoated paper** and *paperboard*, *in rolls or sheets* (Ro. *hârtii şi cartoane necretate în rulouri sau coli*). That is not true with the **printed books**, **newspapers, pictures and other products of the printing industry** (Ro. *cărţi, ziare, ilustrate imprimate şi alte produse ale indus-*

triei poligrafice) there are only two cognate pairs Ro. *rulouri* / Engl. **rolls, imprimat / printed** (without those mentioned earlier).

In the section "*Textiles and textile articles*" as well as in section "*Footwear, headwear, umbrellas*" the translated lexicon (TL) demonstrates that the cognate distance model dominates. It is obvious clearly shown that the terms are of distant origin. Compare: **Carpets and other floor coverings** (Ro. *covoare şi alte acoperitoare* de *podea din materiale textile*), **knitted or crocheted fabrics** (Ro. *materiale textile tricotate sau croşetate*), **articles of apparel and clothing accessorie**s (Ro. *îmbrăcăminte şi accesorii de îmbrăcăminte*), **man-made, staple fibres** (Ro. *fibre sintetice sau artificiale discontinue*), **woven fabrics of synthetic filament yarn** (Ro. *ţesături din fire de filamente sintetice*), **bedlinen, table linen, toilet linen, and kitchen linen** (Ro. *lenjerie de pat, albituri de menaj sau de bucătărie*; (from secţia XII) **footwear, headwear, umbrellas** (Ro. *încălţăminte, pălării, umbrele*), **footwear with outer soles of rubber** (Ro. *încălţăminte cu feţe (talpă din cauciuc)*). There were identified 7 cognate pairs: Ro. *umbrele* / Engl. **umbrellas, textile / textile;** *croşetat* / **crocheted;** *accesorii* / a**ccessories;** *fibre /* **fibres;** *sintetic /* **synthetic;** *toaletă /* **toilet.** The rest pertains to cognate distance.

The products for export trade from section XIII contain articles of stone, plaster, cement or similar materials; ceramic products, glass and glassware, enumerate four items of translation lexicons: articles of asphalt or of similar material (Ro. *articole din asfalt şi din materiale similare*), **ceramic tableware, kitchenware,** other household articles and toilet articles (Ro. *veselă şi alte articole de menaj, de igienă din ceramic*ă), **articles of asbestos-cement**, **of cellulose fibre-cement** (Ro. *articole din azbociment*, **celulozocimet), carboys, bottles, jars, phials (Ro.** *damigene, sticle (garafe), borcane, fiole)*. The following *five cognate pairs* are identified: Ro. *asfalt* / Engl. **asphalt;** *ceramică /* **ceramic;** *azbociment* / **asbestos – cement;** *celulozociment / cellulose (fibre) ciment*; *fiole, ampule* / **phials**. The above-mentioned articles predominately use models of *cognate distance*, as they don't share neither meaning nor similar surface form. Cognate distance models usually arise when both words are derived from different root forms.

The translation lexicon embracing terms such as "Pearls, precious or semi-precious stones, precious metals, metal clads (metale placate) with precious metal, imitation, *jewelry*, *coins* included in section XIV concerning external trade with Russia, Ukraine, Belarus, Romania, Germany enumerate such items as: articles of jewelry and parts thereof (Ro. articole de bijuterie sau de giuvaiergerie şi părţi ale acestora) enumerate more cognate pairs: Ro. perle / Engl. **pearls** ; *meta*l / **meta**l; *preţios, semi-preţios* / **precious, semi-precious;** *imitaţie* / **imitation,** *giuvaiergerie*/ **jewelry**.

The translation lexicons used for rendering articles from base metal, section XV include articles of cast iron, iron, and steel (Ro. *articole din fontă, fier şi oţel*), **stoppers, caps and lids, seals and other packing accessories of base metal** (Ro. *dopuri, capace, sigilii şi alte accesorii* pentru ambalaje din metale comune). The following cognate pairs can be identified: *capace/* **caps**, *accesorii* / **accessories.**

The next sector (XVI) includes a great number of cognate pairs. It covers power and its use. It reflects the electrical machinery and equipment, reproducers, air-conditioning machines, air and vacuum pumps, gas compressors, refrigerators, freezers (congelatoare), washing machines, taps, valves, pipes, tanks (articole de robinetorie), (nivelatoare) as well as transport equipment graders (section XVII). The last sector (XVIII) includes technical novelties such as optical, medical, musical instruments and parts and accessories thereof, where more than 95% are translated using the cognate pairs model. In order to translate arbitrary distant languages, the educators (the translators) use a two-step model via bridge language. (Here the English predominate).

CONCLUSIONS

1. Existing available dictionaries, (an online or a hard-copy one) representing a source language and a language from another language family (Germanic or Slavic) should include intra-family cognate models to yield translation lexicons from the source language to the bridge languages' entire family.

2. The students (the future translators of economic texts) may acquire additional performance of bridge models that can substantially improve language accuracy.

3. Mastering three (at least) or more languages any trainee (student) can identify and use cognate pairs models and cognate distance models in order to overcome contextual non-equivalents.

4. Various learning techniques, and models covering co-training at seminars, could employ these additional measures to create better estimates.

BIBLIOGRAPHY

1. Leon Levițchi, Andrei Bantaș Dicționar englez-român Ed. Teora, București 1995.

2. Gideon S. Mann and David Yarrows Multipath Translation Lexicon Induction via Bridge Language, Baltimore, USA

3. Simard M., G.F. Faster and P. Isabelle Using cognates to align sentences in bilingual corpora Proc. Of TMI – 92, Quebec, 1992.