

ЛИНГВИСТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ АЛГОРИТМА МАШИННОГО ПЕРЕВОДА С РУССКОГО ЯЗЫКА НА ЛИТОВСКИЙ

М. ЗАКАРЬЯН

Алгоритм машинного перевода с русского языка на литовский разрабатывается в Вильнюсском государственном университете им. В. Капсука силами преподавателей и студентов¹. По принципам построения наш алгоритм ближе всего к французско-русскому и венгерско-русскому, о которых сообщалось в печати².

Алгоритм построен на основе сопоставления соотношений грамматических форм и грамматических значений в русском и литовском языках. Сходства и расхождения между двумя языками в указанной области определили структуру алгоритма. Работа по его составлению сводилась, главным образом, к подбору литовских эквивалентов к элементам русского языка. Анализ русского текста и синтез литовского осуществляется в пределах предложения, по предложениям. При этом особенностями литовского языка, его грамматической системы, определяются не только правила синтеза литовского предложения, но и правила анализа русского.

1. Отправным моментом служит учет того обстоятельства, что и в русском и в литовском языках большую грамматическую информацию дает флексия. Поэтому часть правил алгоритма направлена на получение этой информации путем анализа флексий в русском тексте и подбора эквивалентных форм, т. е. форм с теми же грамматическими значениями, в литовском языке. Так как правила алгоритма распределены по частям речи — глагол, существительное, прилагательное и др., то прибегать к анализу других аффиксов (суффиксов, префиксов) приходится редко (например, при анализе русских глаголов исследуются суффиксы прошедшего времени, неопределенной формы и др.)³.

¹ Студенческий кружок прикладной лингвистики: Е. Булите, Л. Винюнайте, А. Гробов, П. Дзикавичюс, А. Евтеев, И. Салтанавичюс, А. Сударикова.

² О. С. Кулагина, Машинный перевод с французского языка, «Известия высших учебных заведений», Математика, 1958, № 5; е е же: О машинном переводе с французского языка на русский, «Проблемы кибернетики», 1960, вып. 4; И. А. Мельчук, О машинном переводе с венгерского языка на русский, «Проблемы кибернетики», 1960, вып. 1.

³ В алгоритме такие суффиксы названы, как и обычные флексии, «окончаниями».

При помощи анализа флексий определяются, например, многие грамматические значения глагола: время, лицо, число, отчасти наклонение. Остановимся на этих правилах.

Мы исходим из того, что указанные грамматические категории в русском и литовском языках совпадают, в основном, по своему грамматическому значению, но различаются, естественно, по формам выражения. Различия в грамматических значениях касаются лишь следующих пунктов: а) в литовском языке в глаголах третьего лица отсутствует значение числа, в то время как в русском оно имеется; б) в простых формах прошедшего времени литовского глагола нет родовых различий, которые имеются в русских (сложные формы прошедшего времени для перевода математического текста не использовались; не использовались также простые формы прошедшего многократного (*būstasis dažninis*) в силу отсутствия соответствующей формы в переводимом тексте); в) в литовском языке в глаголах прошедшего времени и условного наклонения имеется различие в лице, различие, которое отсутствует в русском языке.

Принимая во внимание все сказанное, правила алгоритма были построены на основании следующего соответствия русских и литовских глагольных форм⁴:

Форма	Окончания		Будущее время	
	русские	литовские		
Наст. / буд. вр.				
III л. ед. ч.	-ет -ит	-a -i -o	-s	
I л. мн. ч.	-ем -им	-ame -ime -ome	-sime	
III л. мн. ч.	-ут, -ют -ат, -ят	-a -i -o	-s	
Прошедш. вр.				
ед. ч.	-л -ла -ло	-ai -(i)ai	-o -ė	
мн. ч.			-ote -ėte	
			-o -ė	
Условн. накл.			-tų	
	-л + бы		-tute	
Неопред. форма	-ть -ти -сти	-ti		

⁴ Здесь и в дальнейшем даются только те грамматические формы и их значения, которые имеются в нашем алгоритме, т. е. встречаются в том тексте, на основе которого алгоритм составлен: А. Барсуков, Алгебра, ч.ч. I и II.

Таблица показывает, что русским окончаниям настоящего-будущего времени I лица мн. числа в литовском соответствуют окончания *-ame, -ime, -ome* (в зависимости от типа спряжения) в настоящем времени; окончание *-sime* для будущего (*-s* — формант будущего времени; форма будущего времени в литовском языке образуется от основы неопределенной формы). В третьем лице ед. и мн. числа настоящего-будущего времени русским окончаниям *-ет, -ит, -ут, -ют, -ат, -ят* в литовском в настоящем времени соответствуют окончания *-a, -i, -o*; в будущем же времени к основе неопределенной формы присоединяется формант будущего времени *-s* (личное окончание III лица отсутствует).

Различение настоящего и будущего времени русских глаголов производится в соответствии с видом, указание на который имеется в словаре алгоритма. Таким образом, русские глаголы с пометкой «совершенный вид» переводятся литовскими глаголами будущего времени, а глаголы, не имеющие этой пометки (глаголы несовершенного вида), — литовскими глаголами настоящего времени. Русские глаголы прошедшего времени переводятся литовскими глаголами прошедшего однократного времени (*būstasis kartinis*). Лицо глагола прошедшего времени, как и условного наклонения, определяется по подлежащему — местоимению или другой части речи в им. падеже. Русские глаголы условного наклонения (окончания те же, что и в прошедшем времени, плюс частица *бы*) переводятся литовскими глаголами условного наклонения (*tariamoji nuosaka*) с окончаниями *-tume* для I л. мн. ч. и *-tų* для III л.

Окончаниям русской неопределенной формы *-ть, -ти, -сти* (*-ць* не встречается) в литовском соответствует окончание *-ti*. Функции неопределенной формы глагола в литовском языке (применительно к нашему тексту) в основном совпадают с функциями этой формы в русском, за исключением одного: в русском языке при помощи неопределенной формы образуется форма будущего времени, что не свойственно литовскому. Поэтому русский инфинитив со вспомогательным глаголом (в тексте встречаются лишь формы *будет, будем, будут*) переводится формой будущего времени соответствующего лица и числа, а вспомогательный глагол опускается.

Получить сведения о грамматических значениях русских словоформ по их окончаниям и подобрать соответствующую литовскую словоформу с нужным окончанием удастся и в ряде других случаев: при определении падежа и числа беспредложных существительных, при определении лица и числа личных местоимений и пр. Правила перевода в этих случаях также основаны на таблицах соответствий русских и литовских окончаний. (Местоимения и русского и литовского языков искусственно делятся нами на «основу» и «окончание», например, в русском: *е-го, е-му*; в литовском: *j-is, j-o, j-am* и т. д.)

Но возможности использования окончаний для получения грамматической информации и образования эквивалентов на переводящем языке довольно ограничены, даже для таких языков, как русский и литовский. Удастся это сделать лишь в тех случаях, когда, во-первых, грамматические значения в обоих языках совпадают (во всяком случае, отличаются не по многим пунктам), во-вторых, когда и в русском и в литовском языках данное грамматическое значение выражается только одной формой, если эта форма не омонимична с другой в пределах данной части речи.

II. Как же получается остальная информация о грамматических отношениях в русском тексте и как производится подбор адекватных форм в литовском?

Часть грамматических значений дается непосредственно в словаре. Это — постоянные, неконтекстные (не зависящие от контекста) грамматические значения: часть речи, род существительного в русской и литовской частях словаря. Род существительных, имеющих одно и то же лексическое значение в русском и литовском языках, конечно, часто не совпадает, поэтому указание на род важно для перевода не самих существительных, а согласующихся с ними других слов.

Вид глагола указывается лишь в русской части словаря, так как это помогает отличить глагол будущего времени от настоящего, чего нельзя сделать одним анализом окончаний. В литовском языке вид глагола на образование его временных форм не влияет, поэтому указание на вид не требуется. Указание на возвратность делается в обеих частях словаря, т. к. значения возвратных глаголов в русском и литовском языках не совпадают (хотя способ выражения сходный — русской возвратной частице *-ся, -сь* в литовском соответствует *-si*): литовские глаголы не имеют страдательного значения, и русские страдательные глаголы переводятся страдательными причастиями. Зато литовские возвратные глаголы имеют некоторые грамматические значения, не свойственные русским [ср., например, *pasirinkti* — «выбрать» (себе), *pasiimti* — «взять» (себе, с собою)]. Поэтому соответствующие русские невозвратные глаголы на литовский язык переводятся возвратными.

В русской части словаря указывается также переходность глаголов и деепричастий. Это необходимо для определения падежа управляемого существительного.

Правила перевода русских причастий и деепричастий на литовский язык также основаны, главным образом, на данных словаря: здесь указывается залог, время причастия, вид деепричастия. При составлении правил перевода исходим из следующих русско-литовских соответствий в области причастных форм: двум русским формам — причастию и деепричастию — в литовском соответствуют три: причастие (*dalyvis*), полупричастие (*pusdalyvis*), деепричастие (*padalyvis*). Соотношения меж-

ду русскими и литовскими причастными формами следующие: русским причастиям настоящего и прошедшего времени действительного и страдательного залога в литовском соответствуют те же формы причастия. Кроме того, литовские причастия, как уже указывалось, соответствуют и русским глаголам страдательного залога; если русский глагол стоит в форме прошедшего или будущего времени, в литовском к причастию приписывается связка *buvo* (было) или *bus* (будет). Литовские причастия будущего времени в алгоритме не использовались в силу отсутствия данной категории в русском языке.

Сложнее перевод русских деепричастий: деепричастию совершенного и несовершенного вида в литовском могут соответствовать все три указанные причастные формы. При этом, как показывает анализ нашего текста, соотношения между отдельными формами следующие:

В русск. языке	В литовском языке
деепричастие соверш. вида	{ а) прич. действ. зал. прош. вр. б) дееприч. прош. вр.
деепричастие несов. вида	{ а) дееприч. наст. вр. б) прич. действ. зал. пр. вр. в) полупричастие мн. числа, муж. р.

III. В процессе работы над алгоритмом выяснилось, что хотя в русском и литовском языках окончания дают большую грамматическую информацию, анализ этих окончаний, даже в сочетании с теми сведениями, которые даются непосредственно в словаре, оказывается недостаточным для перевода. Это особенно касается падежа существительных (и других частей речи). Анализ падежного окончания в сочетании с данными словаря о типе изменения⁵ и анализом значения предлога, если он есть, дает сведения о том, в каком падеже стоит русское существительное (или слово, его заменяющее), но не указывает на то, в каком падеже должно стоять переводящее литовское существительное, так как падежные значения в литовском и русском языках, конечно, не всегда совпадают.

Для решения этого вопроса нами были использованы соотношения, существующие между двумя языками в области предложного и беспредложного управления. В тех случаях, когда в русском тексте существи-

⁵ Для сокращения количества омонимичных падежных окончаний, вместо трех русских и пяти литовских типов склонения существительных, в алгоритме различаем 19 русских и 21 (12 — в ед. числе и 9 — во мн. числе) литовских типов изменения по падежам. Для литовского языка — одни и те же типы для существительных и прилагательных. Склонение литовских местоименных прилагательных (*įvardžiutinis būdvardis*) дается как особый тип. Местоимения также распределены по особым типам, параллельно — в русском и литовском языках. Согласно этим типам, к словарю приложены таблицы окончаний.

тельными управляют производные предлоги, перевод на литовский язык, то есть выбор предлога и падежа для переводящего существительного, особых трудностей не представляет, так как хотя управление этих предлогов в двух языках не совпадает, но перевод каждой такой предложно-падежной конструкции в большинстве случаев однозначен. На использование этих соотношений и направлены правила алгоритма. Эти соотношения (во всяком случае, применительно к нашему тексту) оказались следующими:

В русск. языке		В литовск. языке ⁶	
предлог	падеж	предлог	падеж
1 после	P	po	K
2 перед	T	prieš	G
3 впереди	P		
4 навстречу	P		
5 кроме	P	be	K
6 среди	P	tarp	K
7 между	T		
8 вместо	P	vietoj	K
9 согласно	D	pagal	G
10 прежде	P	anksčiau	K

Перевод конструкций с непроизводными предлогами, имеющими также однозначный перевод, аналогичен переводу приведенных выше. Но таких в нашем тексте немного. Это, прежде всего, сочетания предлогов *без, до, из* с родительным падежом в русском языке, которым в литовском соответствуют сочетания *be, iki, iš* также с родительным падежом. Кроме того, сочетанию предлога *под* с творит. падежом в литовском соответствует предлог *po* с творит. падежом, сочетанию предлога *для* с родит. падежом — дательный падеж без предлога.

Остальные русские предложно-падежные сочетания, в зависимости от значения, имеют по два и более эквивалентов в литовском.

По два эквивалента в нашем тексте имеют следующие русские предлоги:

⁶ Приняты следующие сокращения в названиях падежей литовского языка:

Var. — Vardininkas (именительный);

K. — Kilmininkas (родительный);

N. — Naudininkas (дательный);

G. — Galininkas (винительный);

[п. — Įpagininkas (творительный);

V. — Vietininkas (местный).

В русск. языке

В литовск. языке

предлог	падеж	предлог	падеж
1 в	П	— —	V Įn
2 за	Т	už po	K K
3 на	П	iš —	K Įn
4 над	Т	virš su	K Įn
5 о	П.	— apie	G G
6 от	Р	nuo iš	K K
7 по	В	po per	G G
8 при	П	prie —	K Įn
9 с	Т	su —	Įn Įn
10 у	Р	pas —	G K
11 через	В	per —	G Įn

По три эквивалента в литовском имеют следующие русские сочетания:

В русск. языке

В литовск. языке

предлог	падеж	предлог	падеж
1 в	В	i — per	G Įn G
2 за	В	už už per	G K G
3 к	Д	prie i —	K G N
4 по	Д	pagal iš —	G K Įn

Особенно многозначным оказалось сочетание:

на	В	—	[n
		—	K
		—	G
		apie	G
		iš	K
		i	G

Так как значение многозначного предлога (то, каким падежом он в данном случае управляет, а, следовательно, как переводится на литовский язык) зависит от управляющего слова, предлоги эти вводятся в машину не в приведенном виде, а в форме таблицы, где они распределены по группам в зависимости от типа управляющего слова. В каждую такую группу (всего их в нашем алгоритме 30) включаются русские предлоги с соответствующими падежами, управляемые определенными словами. В словаре, под рубрикой «предложный код», слову приписывается номер группы предлогов. Эта помета указывает на то, что данное слово может управлять хотя бы одним из помещенных здесь предлогов с соответствующим падежом. Рядом, в литовской части этой таблицы, указан литовский эквивалент (теперь уже только один). Например, «предложный код» 1 приписан следующим словам: а) *умножить, произведение, умножение, делить, деление, разделить*; они управляют предлогом *на* с винит. падежом. В литовском языке соответствующие слова управляют предлогом *iš* с род. падежом; б) *получится, деление, делится* и др. управляют предлогом *su* с творит. падежом. В литовском — предлогом *su* с творит. падежом и т. д. В таблице под кодом 1 находим:

на В	iš	K
в П	—	V
для Р	—	N
по Д	pagal	G
при П	esant	N
с Т	su	[n

Предложный код 6 имеют, например, слова: *провести, проведенный, проведем*. В таблице под этим кодом имеем:

к Д	i	G
на П	—	V
по Д	pagal	G
от Р	nuo	K
в П	—	V

Таким образом, многозначный предлог может входить в разные группы, но его перевод в каждой такой группе однозначен. Если слово управляет несколькими предлогами, входящими в разные группы, ему приписывается несколько предложных кодов.

В случаях же беспредложного управления наш алгоритм перевода основывается на определенных соотношениях в моделях словосочетаний с беспредложным управлением, существующих между русским и литовским языками. В данной области расхождений немного — литовский язык обнаруживает здесь много черт, сходных с русским: в нем, как и в русском, название объекта при переходных глаголах стоит в винительном падеже. И так же, как в русском, при существительных, образованных от глаголов или соотносительных с ними, название объекта ставится в форме родительного падежа: *составить таблицу* — *sudaryti lentelę*, *составление таблицы* — *lentelės sudarymas*.

Но и в области беспредложного управления есть, конечно, расхождения между этими языками. Как известно, в русском языке имеется ряд глаголов, при которых название прямого объекта ставится не в винительном, а в родительном падеже. Большинство соответствующих литовских глаголов также управляет родительным падежом (например, *vengti* — «избегать», *reikalauti* — «требовать», *ieškoti* — «искать»). Но в некоторых случаях этого совпадения нет: например, русскому словосочетанию *достигать нуля* (род. п.) в литовском соответствует *pasiekti nulį*, где существительное в винительном падеже. По-разному строятся в наших двух языках и словосочетание с прилагательным в сравнительной степени в роли главного слова: беспредложному родительному русского языка в литовском соответствует винит. падеж с предлогом *už* (больше единицы — *didesnis už vieną*).

Для перевода этих беспредложных конструкций была выработана таблица словосочетаний — список типов словосочетаний, встретившихся в нашем тексте, и их литовских эквивалентов:

В русск. языке

В литовск. языке

1 Существ.+существ. в род. падеже	Существ. К+существ.
2 Глагол+существ. в твор. пад.	Глагол+существ. Įn
3 Глагол+существ. в твор. пад.	Глагол+существ. Var
4 Глагол+существ. в твор. пад.	Глагол+существ. G
5 Прилагат.+существ. в род. падеже	Прилаг.+už+существ. G

Номер словосочетания в этой таблице дается в словарной статье управляющего слова.

Правила перевода, основанные на этой таблице и на таблице предлогов, позволяют переводить словосочетания, в которых имеются расхождения между русским и литовским языками, теми же методами, что

и совпадающие, так что факт расхождения между языками не создает дополнительных трудностей при переводе.

Указанные правила алгоритма направлены на анализ окружения переводимого слова. Согласно таким правилам, поиск ведется направо и налево от переводимого слова. В результате такого поиска находим предлог или другое слово, управляющее переводимым словом, и по номеру последнего в таблице определяем падеж переводимого русского слова и падеж переводящего литовского. На анализе окружения построены также правила согласования (прилагательных, причастий и других частей речи).

IV. Некоторые трудности при переводе представляют так называемые «обороты». Это — различные сочетания слов (не обязательно полнозначных), буквальный перевод которых невозможен. При составлении алгоритма эти обороты русского языка были выписаны вместе с их литовскими соответствиями. В нашем тексте их оказалось около 40. Причины их «непереводимости» различны. Часть оборотов — сочетания слов, которым в литовском соответствует одно слово:

иметь в виду — *atsiminti* (букв. «помнить»);
приобрести навык — *įgusti*;
остаться в силе — *galioti*;
принимать во внимание — *atsižvelgti*;
представляет собой — *uga*;
заклЮчить в скобки — *suskliauti*;
заклЮчение в скобки — *suskliautimas*;
привести к общему знаменателю — *(su)bendravidiklinti*;
приведение к общему знаменателю — *bendravidiklinimas*.

Особняком стоит «оборот» *надо* (*было, будет*). В литовском ему соответствует модальный глагол с временными формами: *reikia* (наст. вр.), *reikėjo* (прош. вр.), *reikės* (буд. вр.).

В других случаях причина «непереводимости» оборотов — в семантическом несоответствии составляющих их слов в одном и другом языке. Например, нельзя буквально, «пословно» перевести русское *в самом деле*: его литовский эквивалент *iš tikrųjų* не содержит понятия «дело». То же *надо* сказать и о *ни в коей мере*: литовское *jokiū būdu* буквально означает «никоим образом», *раз навсегда* — *visam laikui* буквально «всему времени», «на все время».

И наконец, в обороты были включены и такие сочетания слов, перевод которых отличается формой главного или подчиненного слова. Они не вошли в таблицу словосочетаний потому, что оказались единичными (в нашем тексте). Вот некоторые из них: *будет видно* — *bus matyti* (букв. «будет видеть»), *другими словами* — *kitaip sakant* («иначе говоря»), *один за другим* — *vienas po kito* («один после другого»), *в порядке записи* — *užrašyto tvarka* («порядком записи»).

Номер оборота в таблице вписывается в словарную статью опорного слова. Таким образом, «оборот» переводится на основе данных словаря и таблицы оборотов.

Итак, информация, необходимая для перевода русского текста на литовский, по правилам алгоритма получается из трех источников: из анализа окончаний, из словаря и приложенных к нему таблиц и из анализа словесного окружения.

Но правила в алгоритме расположены не по признаку источника получаемой информации, а по частям речи. Поэтому в цепи правил перевода одной части речи отдельные правила направлены на получение информации из разных источников: к данным словаря и таблиц прибавляются данные, полученные из анализа окончаний и анализа окружения. Для иллюстрации возьмем 4 правила из цепи правил перевода глагола:

1.0 Если глагол имеет окончание *-ть, -ти, -сти*, переводить инфинитивом (*bendratis*);

1.1 Если влево от инфинитива через один пробел стоит *чтобы* переводить *po gint + инфинитив*;

1.2 Если влево от инфинитива через один пробел стоит вспомогательный глагол *будет, -ем, -ут*, то инфинитив переводить будущим временем (*būsimoji laikas*), а вспомогательный глагол опустить. Лицо и число приписать от вспомогательного глагола;

1.3 Если в положении 1.2 стоит инфинитив *изображаться*, переводить „*bus vaizduojamas*“. Число приписать от русского вспомогательного глагола;

2.0 Если глагол имеет окончание *-л, -ла, -ло, -ли*, искать влево от начала предложения *чтобы* или *бы* или вправо через один пробел *бы*;

2.1 Если есть *чтобы* или *бы*, переводить сослагательным наклонением (*tariamoji nuosaka*);

2.2 Если нет *чтобы* или *бы*, переводить прошедшим временем (*bū-tasis kartinis I.*);

2.3 Если нет *чтобы* или *бы*, а глагол возвратный и в словарной информации есть пометка «св. вид», переводить прошедшим временем (*bū-tasis kartinis I.*);

2.4 Если нет *чтобы* или *бы*, а глагол возвратный и в словарной информации нет пометки «св. вид», переводить *buvo + страдательное причастие настоящего времени (dalyvis, neveik. gūšies esm. I.)*; число приписать от русского глагола;

2.5 Если есть положение 2.1, глагол возвратный и в литовской словарной информации есть пометка „sg“, переводить сослагательным наклонением;

2.6 Если есть положение 2.5, но в литовской словарной информации нет пометки „sg“, переводить *būtų + страдательное причастие настоящего времени*;

2.7 Если есть положение 2.1 и глагол оканчивается на *-ли*, искать слева через один пробел *мы*; если есть, переводить 1-м лицом множественного числа;

2.8 Если есть положение 2.1 и слева от глагола через один пробел стоит *ни*, то к переводящей основе слева приписать „be“ (например, *бы ни были* — *bebūty*);

3.0 Если глагол не имеет окончаний *-ть, -ти, -сти, -л, -ла, -ло, -ли*, переводить настоящим временем, если в словарной информации нет пометки «св. вид»;

3.1 Если в словарной информации есть пометка «св. вид», переводить будущим временем;

3.2 Если глагол 3 лица возвратный и в литовской словарной информации есть пометка „sg“, но в русской словарной информации есть пометка «св. вид», переводить возвратным глаголом будущего времени, 3 лица;

3.3 Если глагол 3 лица возвратный, а в литовской словарной информации нет пометки „sg“ и в русской словарной информации нет пометки «св. вид», переводить страдательным причастием настоящего времени;

3.4 Если глагол 3 лица возвратный, но в литовской словарной информации нет пометки „sg“, а в русской словарной информации есть пометка «св. вид», переводить *bus*+страдательное причастие настоящего времени;

3.5 Если в положении 3.4 стоят глаголы: *потребуется, обратится, повисится, увеличится*, их переводить глаголом будущего времени, 3 лица;

3.6 Если глагол 3 лица возвратный и есть пометка о принадлежности к группе А, переводить глаголом 3 лица настоящего времени;

4.0 Переводящий литовский глагол ставить в том же лице и числе, в котором стоит русский глагол;

4.1 Если глагол прошедшего времени (или условного наклонения), то искать влево через один пробел *я* или *мы*. Если есть — переводить 1 лицом единственного и множественного числа;

4.2 Если *я* или *мы* нет, переводить 3 лицом и т. д.

V. Анализирующие правила нашего алгоритма, построенные на русско-литовских языковых соотношениях, дают возможность не только получить грамматическую информацию о русском слове, но и подобрать литовский эквивалент. Таким образом, на долю синтезирующих правил остается лишь определение способов образования нужной литовской словоформы. Они составлены так, что к полученной из словаря литовской основе приписывается окончание согласно данным, полученным в результате анализа. Делается это при помощи таблиц окончаний, составленных по частям речи. Номер группы окончаний дан в словаре, так что работа по образованию нужной литовской формы сведется к приписыванию к основе соответствующего окончания.

Синтезирующие правила должны определять и порядок расположения слов в переведенном тексте. В нашем алгоритме учтено лишь основное расхождение в порядке слов в русской и литовской фразе — постановка несогласованного определения в родительном падеже перед определяемым словом в литовском языке, тогда как в русском оно

стоит после него. Соответствующее правило алгоритма требует перестановки этих форм (после того, как они переведены). Другие особенности расположения слов в литовском языке, имеющие только стилистическое значение, в нашем тексте не встречались и нами не учитывались.

В настоящее время началось программирование алгоритма машинного перевода с русского языка на литовский.

Vilniaus valstybinis V. Kapsuko v.
universitetas,
Rusų kalbos katedra

[teikta
1962 m. spalio mėn.

MAŠININIO VERTIMO IŠ RUSŲ KALBOS Į LIETUVIŲ KALBĄ ALGORITMO LINGVISTINIAI PAGRINDAI

M. ZAKARJAN

R e z i ū m ė

Straipsnyje aprašomas mašininio vertimo iš rusų kalbos į lietuvių kalbą algoritmas. Jame yra daugiau kaip 1000 žodžių kamienų žodynas, analizės ir sintezės taisyklės.

Žodynas duoda ne tik žodžių vertimą, bet ir kai kurią gramatinę informaciją (nepriklausančią nuo konteksto).

Analizės taisyklės reikalingos priklausančiai nuo konteksto informacijai gauti. Ta informacija gaunama analizuojant žodžių galūnes, žodžių aplinką.

Lietuvių kalbos sakinių sintezė daroma pagal lenteles ir atitinkančias taisykles.
