

Apertium kaj Esperanto

Enkonduko al perregula maŝintradukado al kaj el Esperanto per malfermkoda platformo

Jacob Nordfalk

Kopenhaga Universitata Altlernejo de Inĝenieriko, Danio

jano@ihk.dk

Hèctor Alòs i Font

Ĉuvaŝa Ŝtata Pedagogia Universitato, Rusio

hectoralos@gmail.com

Resumo

Apertium estas rimedaro por konstrui aŭtomatajn tradukilojn. Ĝi estas bazita sur reguloj kaj uzas malprofundan sintaksan analizon. La programaro estas malfermfonta kaj libera, kaj same ĝiaj datumaroj estas liberaj. Ekde ĝia apero en 2005 ĝis februaro 2011 estas publikigitaj 42 tradukiloj inter 27 lingvoparoj, el kiuj el la kataluna, hispana, franca kaj angla al Esperanto kaj el Esperanto al la angla. Ĝia strukturo permesas, ke nefakuloj pri komputila lingvistiko povas evoluigi tradukilojn, kiuj donas kontentigajn rezultojn. Krome, tio, ke Apertium estas malferma, universitate stirata projekto, permesas, ke studentoj evoluigantaj tradukilojn ricevu monhelpojn, ankaŭ se temas pri tradukiloj el kaj al Esperanto. Ĉi tiu artikolo iom detale prezentas Apertium kaj ties Esperanto-tradukilojn kaj ilian funkciadon. Ĝi diskutas la avantaĝojn kaj malavantaĝojn de la kreado de aŭtomataj tradukiloj per Apertium kaj informas pri lingvoj relative facile enkondukeblaj en ĝin.

Ŝlosilvortoj: aŭtomata tradukado, Apertium, Esperanto

1. Kio estas Apertium?

Apertium estas maŝintraduka sistemo bazita sur reguloj, kiu uzas malprofundan sintaksan analizon: t.e. ĝi estas nek statistika nek ekzemplo-bazita kaj ĝi ne faras sintaksan analizon de tutaj frazoj. Ĉiu tradukilo inter du lingvoj funkcias sendepende de la aliaj (ne ekzistas ekz. pivota interlingvo), tamen eblas recikli materialojn de aliaj tradukiloj – kaj oni efektive tion abunde faras. La programaro konsistas en motoro interpretanta datumojn kaj regulojn, en vortaroj kaj en helpilaro por manipuli lingvajn datumojn. Ĝi estas libera kaj malfermfonta, preskaŭ ĉiam laŭ licenco GNU GPL. Tio permesas kunlaboron de unuopuloj, ĉu por la evoluigo de novaj tradukiloj, ĉu por ilia plibonigo. Kaj tio estas des pli farebla, ke la Apertium-arkitekturo permesas, ke ne nur fakuloj pri lingva komputiko kreu tradukilojn, sed ankaŭ homoj, kiuj estas nek lingvistoj nek komputilistoj.¹

¹ Ekzistas sufiĉe ampleksa faka literaturo pri Apertium; vidu: <<http://wiki.apertium.org/wiki/Publications>>. Ĝia ĉefa referenca priskribo estas Forcada (2008) kaj la plej aktuala detala superrigardo kun komparo kun aliaj sistemoj estas

Apertium estas uzebla ĉe la retadreso <<http://www.apertium.org>>² kaj eblas elŝuti la fontkodon por ruligi ĝin en propra maŝino ĉe <<https://sourceforge.net/projects/apertium/>>.

2. Historio de Apertium

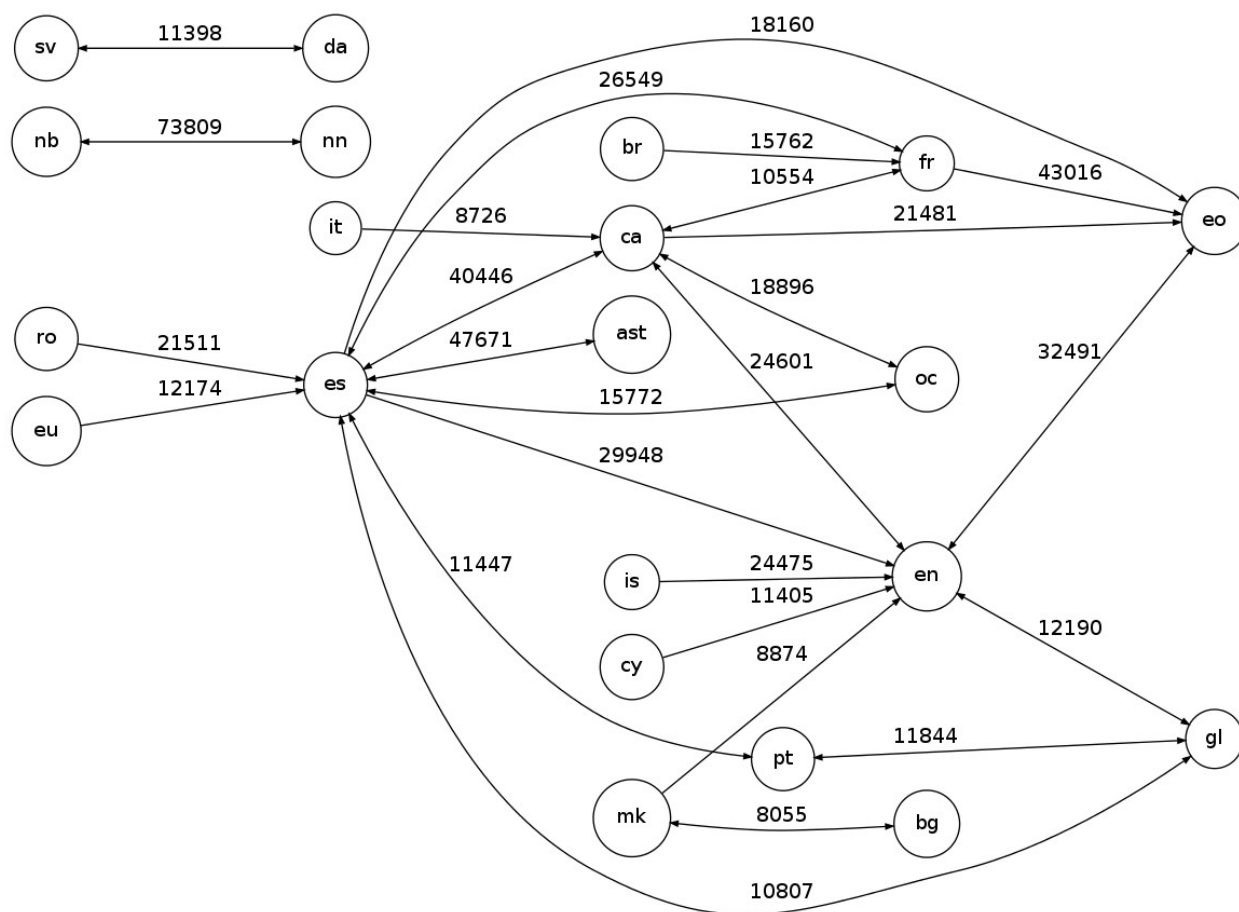
Apertium naskiĝis en 2005 kiel komuna laboro de tri laborgrupoj en Hispanio por krei liberan platformon por aŭtomata tradukado inter la iberiaj latinidaj lingvoj. Tiu apero ne estis hazarda. Kiel konate, Hispanio fariĝis en la lastaj jardekoj unu el la referencaj landoj en la mondo pri plurlingvismo. Kvankam la hispana daŭre estas la sola lingvo oficiala en la tuta ŝtato, ekzistas kvar aliaj regione oficialaj lingvoj (kataluna, galega, eŭska kaj okcitana) kaj kromaj kvar protektitaj (asturi-leona, aragona, hispana signolingvo kaj kataluna signolingvo). Entute pli ol 40% de la hispanianoj loĝas en aŭtonomaj komunumoj kun du aŭ tri oficialaj lingvoj. Sed, plej ofte, ne temas nur pri administra oficialaĵo, sed pri tio, ke la lingvoj estas efektive uzataj en la socio, la kulturo, la administrado, la instrusistemo, la entreprenoj ktp: ekzemple, la plej alte taksata universitato de Hispanio, la universitato de Barcelono, instruas 70%-e en la kataluna lingvo. Entute tiu situacio spronis viglan lingvoindustrion en la lando, i.a. pri aŭtomata tradukado. Tiun forton montras, ekzemple, la publikigo de deko da ĵurnaloj en malsamaj lingvaj eldonoj danke al aŭtomata tradukado.

En la sama jaro 2005, en kiu Apertium estis lanĉita, aperis la tri unuaj tradukiloj inter la hispana, unuflanke, kaj la galega, la kataluna kaj la portugala, aliflanke. En 2007, por evoluigi pli kapablan tradukilon inter la kataluna kaj angla, aperis dua versio de Apertium, kiu donas pli da ebloj por traduki inter (iom) pli malproksimaj lingvoj. Fine de la sama jaro aperis tria versio, kiu ekpermesis la uzadon de Unikodo.

Apertium nuntempe inkluzivas 42 tradukilojn inter 27 lingvoparoj (vd. ilustraĵon 1). Temas pri latinidaj lingvoj (asturia, franca, hispana, itala, kataluna, okcitana, portugala, rumana), ĝermanaj (angla, dana, islanda, norvega bokmål, nov-norvega, sveda), keltaj (bretona, kimra), slavaj (bulgara, makedona), la eŭska lingvo kaj Esperanto.

Forcada (aperonta). Krome la Apertium-vikio donas multajn informojn: <<http://wiki.apertium.org/>>.

² Pro rekonstruo de la TTT-ejo de Apertium, la Esperanto-tradukiloj inter la angla kaj Esperanto kaj el la franca al Esperanto, uzantaj ecojn de la nova versio 3.2 de la Apertium-kerno, nuntempe ne funkcias en tiu TTT-ejo. Ili estas atingeblaj ĉe <<http://apertium.saluton.dk>>. En tiu ĉi adreso, cetere, eblas ruli la aktuale evoluigatajn versiojn de la tradukiloj el kaj al Esperanto.



*Ilustraĵo 1: Publikigitaj tradukiloj de Apertium (februaro 2011).³
La eĝoj montras la nombrojn de vortoj en la dulingvaj vortaroj*

Inter la kurantaj evoluigoj, kvin staras en progresanta stacio (afrikansa-nederlanda, kataluna-sarda, franca-portugala, hispana-itala kaj nord-samea-norvega bokmål) kaj naŭ en komenca (kataluna-rumana, dana-norvega bokmål, angla-Afrikanso, feroa-islanda, irlanda-skota gaela, pola-ukraina, nord-samea-Lule-samea, taĝika-persa, hindia-Urduo). Por multaj pliaj lingvoj estas kolektitaj nur materialoj, ĉefe liberaj vortaroj, kaj eventuale estas ekkonstruita framo de tradukilo, sen efektiva evoluigo de tradukiloj. Inter aliaj, eblas ankaŭ mencii pro siaj plurmilvortaj vortaroj la albana, bengala, kurda, litova kaj slovaka lingvoj.

Rimarkeblas el la lingvo-listo, ke Apertium estas aparte malfermita por t.n. "lingvoj kun malpli da rimedoj". Fondintoj kaj ŝlosilaj evoluigantoj de Apertium aktivas ankaŭ en la asocio SALT MIL, celanta la kreon de elektronikaj rimedoj por minoritataj lingvoj⁴. Por tiaj ĉi lingvoj tradukiloj surbaze de reguloj ŝajnas pli promesplenaj ol la statistikaj, ĉar la statistikaj metodoj postulas tre grandan kvanton da dulingvaj tekstoj, kio ne ekzistas en multaj el ili.⁵

³ Lingvokodoj (laŭ ISO-639 -kodo): ast: asturia; bg: bulgara; br: bretona; ca: kataluna; cy: kimra; da: dana; en: angla; eo: Esperanto; es: hispana; eu: eŭska; fr: franca; gl: galega; is: islanda; it: itala; mk: makedona; nb: norvega bokmål; nn: nov-norvega; oc: okcitana; pt: portugala; ro: rumana; sv: sveda.

⁴ The ISCA Special Interest Group on Speech and Language Technology for Minority Languages <<http://ixa2.si.ehu.es/saltmil/>>.

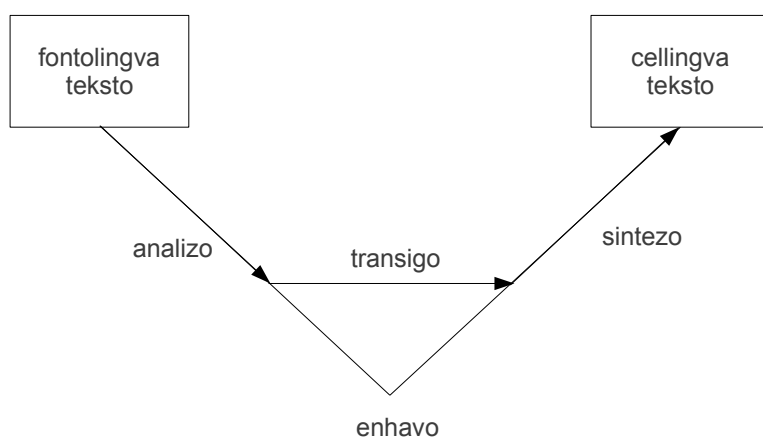
⁵ Rimarku ekz, ke la Lule-samea, por kiu oni evoluigas tradukilon, estas parolata de 1500-2000 homoj: Apertium-agantoj estas ja malfermitaj al kreado de lingvorimedoj por malgrandaj lingvoj.

3. Regul-bazita aŭtomata tradukado

Apertium uzas sufiĉe klasikan traduksistemon, bazitan sur t.n. transigo (angle: *transfer*). La ideo, sur kiu kuŝas tiu paradigmo, estas, ke ofte ne necesas ĝisfunde kompreni tekston, ĉar ĝiaj eventualaj pluraj interpretoj, surbaze ekz. de la larĝeco de la semantikaj kampoj, la plursignifeco de la vortoj aŭ la sintaksa ambigueco, estos samaj en la cellingvo. Plej simple, se ni devas traduki la anglan frazon "My grand-father came" al Esperanto, ni povas diri "Mia avo venis", ne zorgante pri tio, ĉu tiu avo estas la patra aŭ la patrino: tio same malklaras aŭ estas subkomprenata en ambaŭ lingvoj. Simile, frazon kiel "He ate the cookies on the couch" eblas traduki kiel "Li manĝis la kuketojn sur la kanapo", kvankam ne klaras el la frazo mem, ĉu li aŭ la kuketoj kuŝas sur la kanapo. Alivorte esprimite: paŭsado povas sufiĉi por traduki. Kompreneble, ju pli du lingvoj estas lingvistike kaj kulture proksimaj, des pli ĝustos tiu aserto, kaj ju pli lingvistike kaj kulture ili foras, des pli da nekompreneblaĵoj kaj teruraĵoj produktos tiaspeca paŭsado.

Tamen ne malofte, eĉ inter proksimaj lingvoj, necesas iri trans simplan laŭvortan (aŭ laŭvorteran) tradukon. Kiel ni vidos pli detale, eblas enkonduki en la traduk-sistemon regulojn, laŭ kiuj morfologiaj trajtoj aŭ sintaksaj strukturoj aliĝas de la fontolingvo al la cellingvo. Gravas substreki, ke tiuj reguloj estas eksplicite vortumitaj de evoluiganto de la aŭtomata tradukilo, kaj ke ili praktike ĉiam dependas de formaj aspektoj de la lingvo (precipe de sintakso), sed ne de semantiko.

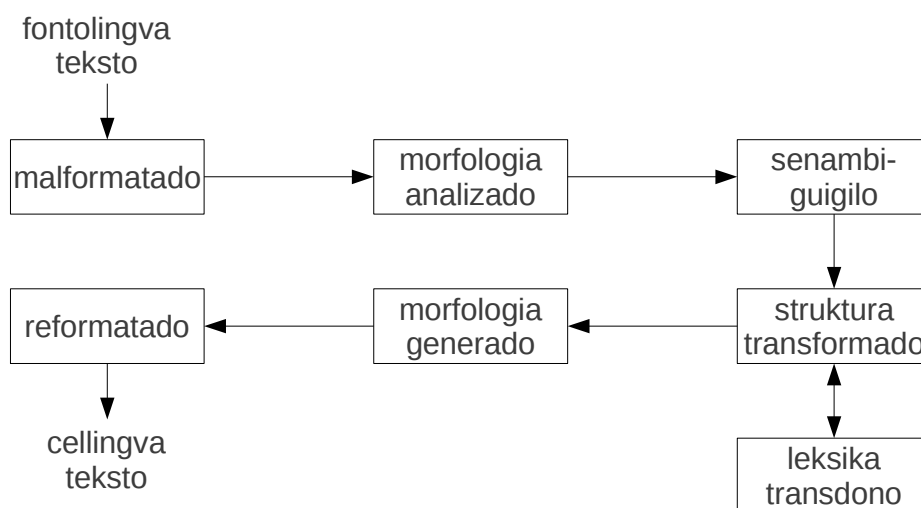
Tiel la procezon de tradukado bazita sur reguloj oni povas skeme prezenti, kiel tio, ke oni unue faras analizon de fontolingva teksto (morfologian, sintaksan), poste transigas tiujn elementojn al formoj kaj strukturoj laŭaj al la cellingvo (ĉu per simpla paŭso, ĉu aplikante transform-regulojn) kaj laste sintezas cellingvan tekston surbaze de tiuj formoj kaj strukturoj (vd. ilustraĵon 2).



Ilustraĵo 2: Procezo de transiga tradukado

4. Kiel funkcias Apertium?

La tradukiloj de Apertium konsistas el apartaj programoj, kiujn oni ĉene kunrulas. Tiel la traduko estas farata en etapo: malformatado, morfologia analizo de la fontlingva teksto, senambiguigo, struktura transformado al la cellingvo, morfologia generado de la cellingva teksto kaj reformatado (vd. ilustraĵon 3). Ĉiu etapo konsistas el unu aŭ pliaj programoj, ĉiu el kiuj ricevas datumojn de la antaŭa, transformas ilin kaj plusendas al la posta.



Ilustraĵo 3: La traduka ĉeno de Apertium

Malformatado konsistas en la flankeniĝo de la informoj pri formato de la originala teksto (ekz. la HTML-kodoj, se temas pri retpaĝo). La *morfologia analizo* okupiĝas pri la serĉo de la *lemo* (kanona formo) de la vortoj kaj ties morfologiaj informoj (t.e. ilia vort-kategorio, nombro, kazo, tenso ktp). Pli-malpli ofte, depende de la lingvo, unu vorto povas esti diverse analizata: ekz. en Esperanto la vorto *dum* povas esti prepozicio aŭ subjunkcio. En tiuj okazoj, la senambiguilo de Apertium decidas, kiu el la eblaj formoj ĝustas, surbaze de la ĉirkaŭaj vortoj. Sekvas la *struktura transformado*, kiu estas la kerna parto de la procezo. Ĝi tradukas la lemojn kaj faras morfologiajn kaj sintaksajn transformojn. El la cellingvaj lemoj kun ties morfologiaj datumoj oni generas la vortojn. Laste la teksto estas *reformatigita* surbaze de la informoj flankenlasitaj en la unua paŝo.

Entute la ĉeno estas facile manipulebla por faciligi ŝanĝojn aŭ aldonojn. Fakte multaj lingvaj paroj ŝanĝas ĝin por siaj bezonoj kaj aldonas ekz. ortografian etapon (ekz. la kunfandon de prepozicio *de* kaj la artikolo *o* al *do* en la portugala), kromajn transigajn nivelojn aŭ asistan senambiguigilon (aŭ eĉ anstataŭigas la defaŭltan de Apertium per alia). La internaj interfacoj inter la programoj estas klare difinitaj tiucele.

5. La Apertium-vortaroj

Esenca aspekto en la arkitekturo de Apertium estas, ke kreantoj de tradukiloj laboras nur super lingvaj datumoj: ili laboras kiel lingvistoj, ne kiel programistoj. La ĉefaj elementoj, kiujn ili prilaboras, estas vortaroj kaj regul-dosieroj.

En ĉiu tradukilo ekzistas precize tri vortaroj: po unu unulingva vortaro por la du koncernataj lingvoj kaj unu dulingva vortaro. La unulingvaj vortaroj havas morfologiajn informojn, kiuj permesas jen rekoni ĉiujn morfologiajn formojn de la vorto (se oni tradukas el tiu lingvo) jen generi la surfacajn formojn surbaze de la morfologia priskribo (se oni tradukas al tiu lingvo). Tiujn du formoj interpreti la vortaron oni konvencie nomas legado *dekstren* aŭ *maldekstren* de la vortaro kaj respondas al la anglalingvaj mallongigoj `<r>` (*right*, dekstre) kaj `<l>` (*left*, maldekstre).

La unulingvaj vortaroj esence estas dividitaj en du partoj. En unu estas deklaratitaj la lemoj. Ĉi tiu deklaro estas kutime triparta: oni atribuas konvencian nomon al la lemo, oni donas ĝian "radikon" (aŭ, pli precize, radikecan formon neŝanĝeblan en ĉiuj morfologiaj formoj de la vorto) kaj oni ligas al tiu ĉi lemo ŝablonon, kiu difinas ĉiujn eblajn vortoformojn. Ĝuste la dua parto de la vortaroj difinas tiujn ŝablonojn. Indas rimarkigi, ke la difino de tiuj ŝablonoj estas kompleksa en morfologie riĉaj fleksiaj lingvoj, kiel la latinidaj. Ekzemple, en la hispana unulingva vortaro de la hispana-esperanta paro, temas pri ĉ. 350 malsamaj ŝablonoj, kies difino okupas 40% pli da spaco ol la pli ol 18 000 vortoj de la lemaro. Pro tio estas multe pli facile krei tradukilon por lingvo jam enkondukita en Apertium, ol mem krei lingvoŝablonaron. Aliflanke, pro la sama kialo, estas Apertium-kunlaborantoj, kiuj kreas unulingvajn vortarojn por faciligi la laboron de tradukil-kreontoj. Tabeloj 1 kaj 2 ekzempligas la difinon de lemoj kaj ŝablonoj en unulingvaj vortaroj kaj tabelo 3 prezentas ekvivalenton de la difino de lemo, aplikante la difinon de ties atribuita ŝablono.

<code><e lm="kato"></code>	<code><i>kat</i><par n="o__n"/></e></code>
<code><e lm="ligo"></code>	<code><i>lig</i><par n="o__n"/></e></code>
<code><e lm="ligi"></code>	<code><i>lig</i><par n="i__vblex"/></e></code>

Tabelo 1: Deklaro de la lemoj katoj, ligo kaj ligi en unulingva Esperanto-vortaro (kun refero al la ŝablonoj `o__n`, respektive `i__vblex`)

<code><pardef n="o__n"></code>			
<code><e><p></code>	<code><l>o</l></code>	<code><r>o<s n="n"/><s n="sg"/><s n="nom"/></r></code>	<code></p></e></code>
<code><e><p></code>	<code><l>oj</l></code>	<code><r>o<s n="n"/><s n="pl"/><s n="nom"/></r></code>	<code></p></e></code>
<code><e><p></code>	<code><l>on</l></code>	<code><r>o<s n="n"/><s n="sg"/><s n="acc"/></r></code>	<code></p></e></code>
<code><e><p></code>	<code><l>ojn</l></code>	<code><r>o<s n="n"/><s n="pl"/><s n="acc"/></r></code>	<code></p></e></code>
<code></pardef></code>			

Tabelo 2: Deklaro de la ŝablono `o__n` en unulingva Esperanto-vortaro⁶

⁶ Kelkaj el la konvencie uzataj gramatikaj etikedoj en Apertium estas: `<n>`: nomo; `<adj>`: adjektivo; `<vblex>`: verbo; `<sg>`: singularo; `<pl>`: pluralo; `<sp>`: singularo aŭ pluralo; `<nom>`: nominativo; `<acc>`: akuzativo; `<gen>`: genitivo. Ekzemploj: `kato<n><pl><nom>` = katoj; `kato<n><sg><acc>` = katon; `kato<adj><pl><acc>` = katajn

<e><p><l>kato</l>	<r>kato<s n="n"/><s n="sg"/><s n="nom"/></r></p></e>
<e><p><l>katoj</l>	<r>kato<s n="n"/><s n="pl"/><s n="nom"/></r></p></e>
<e><p><l>katon</l>	<r>kato<s n="n"/><s n="sg"/><s n="acc"/></r></p></e>
<e><p><l>katojn</l>	<r>kato<s n="n"/><s n="pl"/><s n="acc"/></r></p></e>

Tabelo 3: Ekvivalenta difino de la lemo "kato" surbaze de la antaŭaj difinoj de la lemo kaj ŝablono⁷

Kiel dirite, eblas interpreti tiujn deklarojn analize aŭ sinteze, depende de tio, ĉu oni rekonas fontolingvan tekston aŭ generas cellingvan tekston. Tial, se oni legas la vortaron de maldekstre dekstren, eblas rekoni la vorton "katon", kiun oni anstataŭigos en la interna traduk-procezo per "kato<n><sg><acc>". Male, se en tradukado al Esperanto, alvenas tekstaĵo kiel "kato<n><pl><acc>", oni sintezos la vorton "katojn"⁸.

Eblas difini ankaŭ plurvortajn lemojn, ekzemple sinsekvojn kiel "pensi pri" aŭ "Unuiĝintaj Nacioj", tiel ke ili ricevu specifan tradukon en la dulingva vortaro. Tio ne ĉiam facilas nuntempe en Apertium, ekzemple en plurvortaĵoj, kiuj povas ŝanĝiĝi en pluraj lokoj de la literĉeno (ekz. Unuiĝintaj Nacioj/Unuiĝintajn Naciojn).

Dulingvaj vortaroj estas eĉ pli simplaj: ili principe ligas unu lemon de la fontlingvo al unu lemo de la cellingvo. Kiel tio estas farata, eblas vidi en tabelo 4.

<e><p>	<l>kato<s n="n"/></l>	<r>cat<s n="n"/></r>	</p></e>
--------	-----------------------	----------------------	----------

Tabelo 4: Deklaro de la traduko de "kato" al "cat" (de maldekstre dekstren) kaj de "cat" al "kato" (de dekstre maldekstren)

Ekzistas unusola vortaro uzata por ambaŭ traduk-direktoj. Denove por tio oni uzas la konvencion "legi" ilin de dekstre maldekstren aŭ inverse. En tiu okazo, ekzemple, se oni legas la vortaron de maldekstre dekstren oni havas la tradukon el Esperanto al la angla. Tiel, se en la procezo de tradukado atingas tiun vortaron la tekstaĵo "kato<n><pl><acc>", ĝi estos "tradukita" al "cat<n><pl><acc>" per simpla anstataŭigo de signoj. Postaj paŝoj en la tradukprocezo devos forigi la informon "<acc>", fremdan al la morfologio de la angla lingvo, kaj la rezulta tekstaĵo "cat<n><pl>" estos sintezita al "cats", per la procedo vidita ĉi-supre.

Esenca limigo de la dulingvaj vortaroj estas, ke por fontolingva lemo ekzistu nur unusola traduko, sed cellingva lemo povas traduki plurajn fontolingvaj (vidu ekzemplon en tabelo 5). Eblas ankaŭ esprimi malsamajn tradukojn de lemo depende de morfologiaj ecoj, kiel genro aŭ nombro. Tiel ekzemple singulara "ĉiu" estas tradukata al la angla "every", sed plurala "ĉiuj" fariĝas "all" (ĉiu studento → every student, ĉiuj studentoj → all students).

<e>	<p><l>belorusa<s n="adj"/></l>
	<r>Belarusian<s n="adj"/></r></p></e>
<e r="RL">	<p><l>belorusa<s n="adj"/></l>
	<r>Byelorussian<s n="adj"/></r></p></e>

⁷ La indiko *lm*="kato" de la XML-elemento <e/> estas flankenlasita, ĉar ĝi estas nura krominformo.

⁸ En la vortara formato la XML-elemento <s/> reprezentas etiketon, tiel ke ekzemple <s n="acc"> reprezentas la etiketon <acc>.

```
<e r="LR"><p><l>bjelorusa<s n="adj"/></l>
<r>Belarusian<s n="adj"/></r></p></e>
```

Tabelo 5: Ekzemplo de kompleksa deklaro laŭ la direkto inter la angla kaj Esperanto:
belorusa ↔ Belarusian; bjelorusa → Belarusian; belorusa ← Byelorussian

Ĉiuokaze, eĉ se Apertium permesas kelkajn flankajn vojojn (ankaŭ per la starigo de reguloj), la limigo de traduko de unu lemo al unusola lemo sentiĝas kiel grava limo en la kvalito de la tradukoj de Apertium. Ekzemple, apenaŭ eblas nuntempe decidi laŭ la kunteksto, ĉu la anglan vorton "table" oni traduku per "tablo" aŭ "tabelo" aŭ ĉu "accent" fariĝu "akcento", "akĉento" aŭ "supersigno". Tamen fojfoje sintakso povas helpi, kaj Apertium povas ekspluati tion: ekzemple principe la franca vorto "place" estas tradukita kiel "loko", tamen eblas difini en la vortaro, ke ĝi devas esti "placo" antaŭ la prepozicio "de": tio solvas ne malgrandan parton de la mistradukoj. Entute tiu grava malfacilaĵo kaŭzas foje tre ridindajn erarojn, kiel tio, ke en la kataluna-esperanta tradukilo "invasió dels gots" fariĝas "invado de la glasoj" anstataŭ "invado de la gotoj". Ekzistas ideoj por parte solvi la problemon kaj kelkaj estas elprovataj en tradukiloj⁹, sed ili estas konsiderataj ankoraŭ eksperimentaj kaj ne enkondukitaj kiel norma procedo en la Apertium-tradukiloj.

6. Transigaj reguloj

La dua kaj plej esenca flanko en la evoluigo de Apertium-tradukilo estas la verkado de transigaj reguloj. Ĉi tiuj permesas efektivigi, ke la traduko ne estu nur laŭvorta, danke al struktura transformado de la fontolingvo en la cellingvon. La unua versio de Apertium permesis nur unu nivelon de struktura transformado, kio montriĝis kontentiga inter la iberiaj latinidaj lingvoj, sed en la dua versio oni kreis tri sinsekvajn strukturajn transformojn, kio sufiĉe plipotencigas la motoron. Nuntempe eblas uzi eĉ tiom da niveloj, kiom oni deziras: la laste apertinta tradukilo el la franca al Esperanto uzas ok nivelojn.

La transformoj estas difinitaj per t.n. "transigaj reguloj", permesantaj ŝanĝi ecojn kaj strukturojn inter la fonta kaj cела lingvoj (vd. tabelon 6 por ekzemplo de strukturaj ŝanĝoj).

(angla) intelligent pupils → (Esperanto) inteligentaj lernantoj <i>akordigo de nombro por adjektivo kaj nomo</i>
(Esperanto) li manĝas → (angla) he eats <i>akordigo de verbo kaj subjekto</i>
(angla) computer textbook → (Esperanto) komputila lernolibro <i>ŝanĝo de gramatika kategorio: computer (nomo) → komputila (adjektivo)</i>
(angla) John's book → (Esperanto) libro de John <i>ordoŝanĝo kaj aperigo de prepozicio</i>
(hispana) la casa grande y antigua → (Esperanto) la granda kaj malnova domo <i>ŝanĝo de la loko de la adjektivoj rilate al la nomo (casa = domo)</i>
(franca) j'aime ce livre → (Esperanto) mi ŝatas tiun libron

⁹ Vidu Wiechetek (2010).

<i>aldono de akuzativo</i>
(kataluna) treballo → (Esperanto) mi laboras <i>aldono de persona pronomo</i>
(franca) je viens de parler avec lui → (Esperanto) mi ĵus parolis kun li <i>anstataŭigo de la strukturo "venir de verbo-en-infinitivo" per "ĵus verbo-konjugaciita"</i>
(kataluna) no m'agrada aquest llibre → (Esperanto) mi ne ŝatas ĉi tiun libron <i>ŝanĝo de la strukturo (el la tipo 'plaças al mi io', kutima fontolingve, al la pli ofta cellingve 'mi ŝatas ion')</i>
(franca) je veux que tu viennes → (Esperanto) mi volas ke vi venu <i>ŝanĝo de la tenso en la subfrazo, lige al volverbo en la ĉeffrazo</i>
(franca) je disais qu'il viendrait → (Esperanto) mi diris ke li venos <i>ŝanĝo de la tenso en la subfrazo en nereкта parolo</i>

Tabelo 6: Ekzemploj de strukturaj ŝanĝoj (morfologiaj aŭ sintaksaj)

Tuj sekve ni prezentas la tri normajn etapojn, per kiuj la Apertium-tradukiloj dividas la struktur-ŝanĝan taskon.

La unua etapo grupigas la frazelementojn, t.e. surbaze de la morfologie etikedita fontolingva teksto, ĝi unuigas lemojn en frazerojn (ekz. en nom-sintagmojn) kaj aldonas al ĉi tiuj ĉapon kun sintaksaj informoj. Ĝi povas forigi aŭ aldoni elementojn: ekzemple la nedifinita artikolo de la latinidaj kaj ĝermanaj lingvoj estas kutime flankenlasita, kiam oni tradukas al Esperanto, dum la angla genitivo postulas ofte aldonon de la prepozicio "de" en la Esperanto-traduko. En tiu etapo oni ŝanĝas, aldonas aŭ forigas morfologiajn informojn de la fontolingvo por ekhavi cellingvajn: ekzemple la genro de la latinidaj lingvoj plej ofte malaperas en la Esperanto-traduko (kvankam pro facileco ĝi estas en kelkaj okazoj konvencie konservata), dum aperas indiko pri kazo en la nomoj; krome la multaj tensoj kaj modoj de tiuj lingvoj estas simpligataj en Esperanton. Ne ĉiam oni scias en tiu nivelo ĉiujn cellingvajn informojn: ekzemple la kazon oni kutime decidas en la dua etapo.

Ĉiuj ĉi modifoj, aldonoj aŭ forigoj estas difinitaj per transigaj reguloj. Temas fakte pri tre malgrandaj programetoj, koditaj en specifa XML-sintakso, en kiuj oni difinas, kiel transformiĝas la enira tekstoĉeno, se ekzistas difinitaj kondiĉoj (ekz. se estas sinsekvo de nomo kaj adjektivo). La Apertium-tradukiloj kutime havas de 100 ĝis 300 tiaj reguloj en la unua transform-etapo.

En tiu ĉi etapo okazas la traduko de la lemoj. Ĝi estas la sola el la tri normaj, en kiu eblas konsulti la fontolingvajn datumojn (kiel la lemon kaj morfologiaĵojn). Tiuj informoj malaperas, krom se oni gardas ilin en la ĉapo de la kreataj grup-elementoj. Al tiuj ĉi ĉapoj aliras la sekvaj transform-etapoj. La traduko estas principe laŭvorta, tiel kiel estas difinite en la dulingva vortaro, tamen eblas krei (kaj oni efektive kreas) regulojn por elekti aliajn tradukojn, ol la vortare difinitajn, surbaze de la ĉirkaŭaj vortoj aŭ vortotipoj. Tabelo 7 montras ekzemplan de pakado de pluraj vortoj en unu grupon en tiu unua transiga etapo.

Regulo	<det>, <adj>, <n> → grupo <SN>
Efektivigo	¹⁰ \$ ^a<det><ind><sg>\$ ^white<adj><sint>\$ ^cat<n><sg>\$

→ ^det_adj_nom<SN><sg><nom>{^blanka<adj><2><3>\$ ^kato<n><2><3>\$}\$ ¹¹
--

Tabelo 7: Ekzemplo de unuaetapa regulo kaj ties efektivigo en ekzempla frazo

Eblas rimarki, ke Apertium ne faras plenan sintaksan analizon de la frazoj por traduki, sed limiĝas al malprofunda, parta trarigardo. Tiel, la ekzempla regulo aktiviĝos, se dum la inspektado de la frazo laŭ la legodirekto la tradukilo trafas determinilon, sekvitan de adjektivo kaj nomo: pliajn fraz-elementojn oni ĉi tiam neglektas. Povas okazi, ke pluraj reguloj konkurencas: ekzemple en difinita momento eblus lanĉi regulon por sinsekvo "determinilo, adjektivo, nomo" kaj sinsekvo "determinilo, adjektivo, nomo, adjektivo". Estas precize difinite, kiu regulo estas elektita inter konkurencantoj: la plej longa trafo venas unue en la listo de prioritatoj, tiel ke kvar-elementa trafo gajnos kontraŭ tri-elementa.

La dua etapo de transigo konsistas en la apliko de reguloj sur la grupoj starigitaj en la unua. Tio igas, ke surbaze de la sintakse etikedita cellingva teksto de la unua etapo, oni modifas informojn de la ĉapoj, aldonas aŭ forigas elementojn aŭ rearanĝas tutajn frazpartojn. Ekzemple, eblas konstrui regulon similan al tiu ĉi:

$$\langle \text{SN} \rangle \langle \text{nom} \rangle, \langle \text{SV} \rangle, \langle \text{SN} \rangle \langle \text{nom} \rangle \rightarrow \langle \text{SN} \rangle \langle \text{nom} \rangle, \langle \text{SV} \rangle, \langle \text{SN} \rangle \langle \text{acc} \rangle$$

Alivorte: se oni renkontas nom-grupon en nominativo, kiun sekvas verbo, kiun siavice sekvas nom-grupo en nominativo, ŝanĝu la kazon de la dua nom-grupo al akuzativo. Kompreneble tia regulo postulas multe da polurado, sed ĝi baze respegulas kiel eblas aldoni akuzativon en traduko al Esperanto el senkaza lingvo kun ordo subjekto-verbo-objekto, kiel (praktike) la latinidaj aŭ la angla.

Ĉi-okaze sufiĉus ŝanĝi la informojn de la grup-ĉapoj kreitaj en la unua etapo, sed foje necesas fari pli kompleksajn operaciojn. Ekzemple en la hispana plej ofte subjektaj pronomoj ne estas uzataj, sed oni subkomprenas ilin el la kunteksto. Do en tiu etapo necesas rimarki, ke subjekto mankas en tia frazo kiel "Como" kaj aldoni ĝin por ricevi "Mi manĝas", ĉar tiu pronoma subjekto estas postulata en Esperanto.

La tria kaj lasta etapo de la transiga procezo estas la plej simpla. Ĝi esence konsistas nur en la elpako de la grupoj, heredigante la informojn de la ĉapo al la leksikaj elementoj kaj flankenlasante la frazer-ĉapon. Tiel ekzemple

$$\langle \text{SN} \rangle \langle \text{sg} \rangle \langle \text{acc} \rangle \{^{\wedge} \text{kato} \langle \text{n} \rangle \langle \text{2} \rangle \langle \text{3} \rangle \$\}$$

fariĝus

$$\text{kato} \langle \text{n} \rangle \langle \text{sg} \rangle \langle \text{acc} \rangle$$

Tabelo 8 montras komentitan laŭpaŝan ekzemplan de traduk-procezo.

¹⁰ Unuopaj leksikaj unuoj situas inter cirkumflekso (^) kaj dolar-signo (\$).

¹¹ La ĉapo situas antaŭ la frazero, kiu estas ĉirkaŭita de vinkuloj { kaj }. Inter la vinkuloj la morfologiaj informoj de la leksikaj unuoj povas havi numerojn celantajn la tioman elementon inter triangulaj krampoj de la ĉapo (ekz. ĉi-ekzemple <2> ŝanĝas al <sg> kaj <3> al <nom>). Tiel oni agas en Apertium, ĉar en la niveloj post la unua, oni povas manipuli nur la ĉapojn de la grupoj.

Etapo	Teksto fine de ĉiu etapo	Rimarkoj
	I saw the cat's tail	
Morfologia analizado	<code>^I/I<num><sg>/prpers<prn><subj><p1><mf><sg>\$</code> <code>^saw/saw<n><sg>/see<vblex><past>\$</code> <code>^the/the<det><def><sp>\$</code> <code>^cat/cat<n><sg>\$^'s/'s<gen>\$</code> <code>^tail/tail<n><sg>\$</code>	"I" povas esti kaj romia cifero (=1) aŭ persona pronomo (=mi). "saw" povas esti nomo (=segilo) aŭ verbo (=vidis).
Senam-biguigo	<code>^prpers<prn><subj><p1><mf><sg>\$</code> <code>^see<vblex><past>\$</code> <code>^the<det><def><sp>\$</code> <code>^cat<n><sg>\$^'s<gen>\$</code> <code>^tail<n><sg>\$</code>	Oni decidas, surbaze de la morfologio de la ĉirkaŭaj vortoj, ke "I" estas pronomo kaj "saw" estas verbo.
1a transiga etapo	<code>^prnpers<SN><sg><nom>{</code> <code> ^prpers<prn><p1><mf><sg><3>\$</code> <code>}\$</code> <code>^verb_all<SV><aliaj><past>{</code> <code> ^vidi<vblex><3>\$</code> <code>}\$</code> <code>^det_nom_gen_det_nom<SN><sg><nom>{</code> <code> ^la<det><def><sp>\$</code> <code> ^vosto<n><sg><3>\$</code> <code> ^de<pr>\$</code> <code> ^la<det><def><sg><nom>\$</code> <code> ^kato<n><sg><nom>\$</code> <code>}\$</code>	La frazo estas analizita kiel nom-sintagmo (SN=mi), verb-sintagmo (SV=vidis), nom-sintagmo (SN=la vosto de la kato). La kazo de "vosto" dependas de tiu de la ĉapo (kaj provizore ĝi estas difinita kiel nominativa), dum tiu de "kato" estas senrevene fiksita kiel nominativa.
2a transiga etapo	<code>^prnpers<SN><sg><nom>{</code> <code> ^prpers<prn><p1><mf><sg><3>\$</code> <code>}\$</code> <code>^verb_all<SV><aliaj><past>{</code> <code> ^vidi<vblex><3>\$</code> <code>}\$</code> <code>^det_nom_gen_det_nom<SN><sg><acc>{</code> <code> ^la<det><def><sp></code> <code> ^vosto<n><sg><3>\$</code> <code> ^de<pr>\$</code> <code> ^la<det><def><sg><nom>\$</code> <code> ^kato<n><sg><nom>\$</code> <code>}\$</code>	En la ĉapo de la tria frazparto "nom" ŝanĝiĝas al "acc",
3a transiga etapo	<code>^prpers<prn><p1><mf><sg><nom>\$</code> <code>^vidi<vblex><past>\$</code> <code>^la<det><def><sp>\$</code> <code>^vosto<n><sg><acc>\$</code> <code>^de<pr>\$</code> <code>^la<det><def><sg><nom>\$</code> <code>^kato<n><sg><nom>\$</code>	Oni heredigas la morfologiajn informojn de la ĉapo al la interno de la grupoj kaj forigas la ĉapojn.
sintezado	mi vidis la voston de la kato	Oni kreas la "surfacan formon" de la vortoj surbaze de la lemoj kaj la morfologiaj informoj

Tabelo 8: Ekzemplo de traduko el la angla al Esperanto

7. Semantiko

La traduko de tekstoj surbaze preskaŭ ekskluzive de morfologio kaj sintakso estas en kelkaj okazoj malsufiĉa, precipe ĉe plursignifaj vortoj, kiam tiu plursignifeco ne estas egala en la ĉela kaj fonta lingvoj. Sed ekzistas malpli evidentaj okazoj, kiuj povas pli bone prilumi la aspektojn de la problemo.

Ekzemple, en la menciita frazo "Como" de la hispana, eblas kompreni el la verbo, ke la subjekto estas unua persono de singularo, sekve senhezite redonebla per "mi" en Esperanto: "Mi manĝas". Tamen en la frazo "Come", kvankam ni same per la verbo komprenas, ke la subjekto estas la tria persono de singularo, ni ne povas eltiri el la hispana morfologio kaj sintakso, ĉu ni devas uzi en Esperanto "li", "ŝi" aŭ "ĝi" kiel subjekto de "manĝas".

La franca estas latinida lingvo, sekve proksime parenca al la hispana, tamen en la franca la subjekta pronomo estas deviga, malkiel en la hispana. Tio signifas, ke ni alfrontas frazojn pli eksplikitajn, kiel "Il mange", kie "il" estas la vira formo de subjekta singulara pronomo de la 3a persono. Ĉu tio signifas, ke "Il mange" devas esti tradukata kiel "Li manĝas"? Bedaŭrinde ne ĉiam, ĉar "il" povas esti same homo kiel besto, eĉ estaĵo, kaj sekve povas esti tradukenda en Esperanton kiel "li" aŭ "ĝi". Simile, kiam en la portugala aŭ la kataluna oni renkontas la posedan adjektivon "seu" necesas kompreni, kiu estas la referencato por ĝuste traduki ĝin al "lia", "ŝia", "ĝia", "ilia" aŭ "sia": morfologio kaj sintakso ne sufiĉas.

Antaŭ tiaj ĉi problemoj la nuna evoluigo de Apertium apenaŭ donas rimedojn. Tamen substrekindas, ke ekzistas parta solvo de la problemo tre facile enkondukebla en Apertium, se informoj ekzistus. Ekzemple homo tradukus la francan frazon "Elle est bordée par la mer du Nord" kiel "Ĝi estas bordita de la Norda Maro" (aŭ simile), t.e. sen kunteksto klaras, ke nur lokon bordas io. Tiajn informojn, ĉu subjekto aŭ verba komplemento povas esti homa aŭ ne-homa, eblus havi pri la Esperanto-verboj, same kiel ili ekzistas por aliaj lingvoj. Simile multaj tiaj lingvaj informoj estas uzindaj por levi la kvaliton de aŭtomataj tradukoj. Gravas, do, instigi la lingvistikan esploradon de Esperanto kaj kunordigi la diversajn esploristojn, por certigi, ke ĉies atingoj estas konataj aŭ facile troveblaj de aliaj.¹²

8. Esperanto kaj Apertium

Esperanto aperis relative frue en Apertium. Komence de 2007 esplorgupo pri aplikita lingvistiko de la Universitato Pompeu Fabra de Barcelono, kunlabore kun Kataluna Esperanto-Asocio kaj financite de la entrepreno ABC-Enciklopedioj, entreprenis la kreon de du aŭtomataj tradukiloj, el la

¹² Menciindas ankaŭ la problemo de la generado de artikoloj, kiam ili mankas en la fontolingvo, ekz. la generado de nedifinita artikolo, kiam oni tradukas al latinida aŭ ĝermana lingvo el Esperanto, aŭ la generado de difinita artikolo en Esperanto, kiam oni tradukas al ĝi el slava lingvo kiel la rusa aŭ la pola. En tiuj okazoj ŝajnas plej promesplena miksa uzo de statistikaj metodoj al la regul-bazita malferma framo de Apertium. Tamen, kiel en la supre diskutita problemo pri elekto inter malsamaj tradukoj de plursignifa vorto, tiaj aliroj estas ankoraŭ eksperimentaj en la Apertium-komunumo.

kataluna kaj hispana lingvoj, al Esperanto. Unua eldono de tiuj tradukiloj aperis en januaro 2008. En tiu sama jaro Jacob Nordfalk evoluigis tradukilon el la angla al Esperanto kaj sekvajare el Esperanto al la angla. En tiu sama 2009, Hèctor Alòs i Font pretigis novajn versiojn de la tradukiloj el la kataluna kaj la hispana. En februaro 2011 aperis tradukilo el la franca al Esperanto, evoluigita de Hèctor Alòs i Font kun la kunlaboro de Corinne Daudier por la kreado de la unulingva franca vortaro.

Temas, do, entute pri tradukiloj el tri latinidaj lingvoj (kataluna, hispana kaj franca) kaj el la angla al Esperanto, kaj el Esperanto al la angla. Pro sia strukturo Esperanton oni povas konsideri hindeŭropa lingvo, sed ĝi sidas aparte apud la latinidaj, la ĝermanaj, la slavaj aŭ aliaj subgrupoj de hindeŭropaj lingvoj. Tiel la komplekseco de la evoluigo de aŭtomata tradukilo el aŭ al Esperanto montriĝas simila al tiu inter du hindeŭropaj lingvoj de malsamaj subgrupoj. La reguleco de Esperanto efektive tre simpligas ties unulingvan vortaron, ĉar la nombro de necesaj paradigmoj estas tre malgranda kompare al la nombro bezonata por aliaj hindeŭropaj lingvoj, se kompari ties fleksiajn variecojn. Aldone, la morfologia ambigueco de Esperanto estas tre malgranda kompare kun la aliaj lingvoj. Tamen, rilate al la leksika ambigueco ligita al plursignifeco de vortoj, ne klare evidentiĝas pli granda facileco en la direkto el (nacia) hindeŭropa lingvo al Esperanto ol inter, ekzemple, latinida lingvo kaj la angla. Samtempe, montriĝas, ke necesas simila nombro de transigaj reguloj por traduki el (nacia) hindeŭropa lingvo al Esperanto kiel, ekzemple, el latinida lingvo kaj la angla aŭ inverse por havi similajn kvalit-rezultojn. Male, malpli da reguloj sufiĉas por doni similajn aŭ eĉ pli bonajn rezultojn inter la latinidaj lingvoj – kaj certe multe pli bonaj inter la du variantoj de la norvega.

9. Ŝlosilaj aspektoj de Apertium por esperantistoj

Apertium prezentas kelkajn distingajn trajtojn por la esperantistoj, trans teknikaj detaloj, kiuj estas pritaksindaj.

Verŝajne la plej grava aspekto estas, ke temas pri malfermita platformo, universitate stirata, uzata de aro da homoj kaj grupoj en diversaj landoj. Sekve la platformo evoluadas kaj pliboniĝas sendepende de la fortoj de la Esperanto-komunumo. Ankaŭ ĝia videbleco ne dependas de la Esperanto-komunumo aŭ la eblo publikigi fakartikolon pri ĝi, ricevi esplor-stipendion ktp. Tamen ja la Esperanto-parolantoj povas profiti el la atingoj de la Apertium-komunumo kaj kontribui al ĝi.

Elstara ekzemplo de tiu ĉi forta punkto de Apertium estas tio, ke ĝi sukcesis akiri en la lastaj tri jaroj (2009, 2010 kaj 2011) averaĝe po naŭ stipendiojn de Google.¹³ Temas pri stipendioj de po ĉ. 5000 US\$ por universitataj studentoj evoluigontaj novajn kapablojn de Apertium: ĉu kernajn, ĉu flankprogramajn, ĉu novajn tradukilojn. Bedaŭrinde, malgraŭ niaj klopodoj, nek en 2009 aŭ 2010,

¹³ <<http://code.google.com/soc/>>.

nek en 2011 sin proponis studento por krei novan Esperanto-tradukilon, kvankam, en la nuna stato de la aferoj, po unu jare certe estus akceptita.

Aliflanke, la malfermeco de la kodo kaj de la lingvaj rimedoj permesas la kunlaboron de unuopaj evoluigantoj same kiel de malsamaj grupoj. Tio same validas por unuopaj Esperanto-evoluigantoj kiel inter tiuj ĉi kaj evoluigantoj de aliaj lingvoparoj. Krome, tiu eco garantias, ke la uzo de la Apertium-tradukiloj kaj de la lingvorimedoj evoluigataj en ĝi, kiel vortaroj,¹⁴ estas kaj estos universale alirebla kaj senpaga – ankaŭ sendepende de la kvanto da teksto, kiun oni deziras traduki.

Entute Apertium montriĝas relative facile uzebla, tiel ke eblas partopreni en la evoluigo de Esperanto-tradukilo sen grandaj scioj pri komputiko kaj komputika lingvistiko (ĉefe, se staras apude pli fakaj aŭ spertaj gvidantoj). Kun la nuna sperto, niatakse la kreo de baza aŭtomata tradukilo el latinida lingvo al Esperanto (aŭ inversen) postulas 4-6 monato-homojn (600-1000 laborhorojn), depende de la sperto de la evoluiganto kaj de la preteco de la lingvorimedoj; verŝajne tio estos simila por ĝermanaj kaj slavaj lingvoj, se ili jam estas enkondukitaj en Apertium. Per "baza" ni komprenas "tradukanta neliteraturajn tekstojn (ne solajn frazojn) al komprenebla lingvo sen eraroj en ĉiu frazo".

10. Apertium kompare kun aliaj Esperanto-tradukilaj sistemoj

La nuntempe evoluigitaj Apertium-tradukiloj ne estas la plej bonaj inter la ekzistantaj Esperanto-tradukiloj: la angla-esperanta tradukilo de la firmao Gramtrans¹⁵, kreitaj de Eckhard Bick, donas senkonteste pli bonajn rezultojn ol iu ajn rivalo. La kialoj estas analizendaj kaj tre verŝajne pluraj. La tradukiloj de Gramtrans funkcias surbaze de multaj jaroj da profesia tre detala studo de la angla lingvo, kiu permesas havi en la internaj vortaroj multajn pliajn semantikajn kaj sintaksajn informojn, ol ĝis nun la kunlaborantoj de Apertium sukcesis kolekti inter la libere haveblaj informoj. Tiu pli profunda, komerc-cela traesploro de fontolingvo permesas, krom multe pli bona, komerc-nivela senambiguigado, ankaŭ ĝisfundan sintaksan analizon, kiu montriĝas pli sukcesa ol la surfaca frazer-analizo praktikata de Apertium. Laste, ekzistas lertaj arkitekturaj kapabloj en Gramtrans, kiuj superas la nuntempe ekzistantajn en Apertium, eble plej elstare la apartaj reguloj por elekti inter pluraj traduk-ekvivalentoj de fontovorto. Fakte, tiu ĉi kapablo estis lastatempe enkondukitaj en kelkaj tradukiloj de Apertium, sekve ni esperas enkonduki ĝin en ankaŭ la Esperanto-tradukiloj.

Entute, la tradukiloj de Gramtrans estas verŝajne kvalite nesupereblaj de Apertium aŭ de iu ajn grupo de amatoroj libervole kaj libertempe laborantaj. Cetere, provi konkuri ŝajnas tre malsaĝa tempoperdo: oni prefere okupiĝu pri aliaj lingvoparoj netraktotaj de Gramtrans.

¹⁴ Ekzemple, el la kataluna-esperanta vortaro de Apertium, estis eldonita la vortaro plej ampleksa ĝis nun inter tiuj du lingvoj (Alòs 2010). Cetere, rimarkindas, ke la ŝanĝo de formato, por ke la vortaro estu konsultebla kiel normala papera vortaro (en formato PDF) postulis ne pli ol kelkajn horojn da laboro.

¹⁵ <<http://gramtrans.com/>>.

Aliflanke ankaŭ Apertium havas esencajn atutojn: ĝi estas malfermkoda kaj senpage kaj senkondiĉe uzebla. Gramtrans permesas senpage traduki nur 350 vortojn tage, kaj nur por persona uzo sen publikigo. Se oni volas publikigi tekstojn tradukitajn per Gramtrans oni devas pagi 210 € por duonjara abono¹⁶. Aldone, Apertium ne dependas de unusola homo aŭ firmao, sed iu ajn povas plibonigi la ekzistantajn tradukilojn aŭ krei novajn por neekzistanta lingvoparo.

Tabelo 9 montras specimenojn de tradukoj de angla al Esperanto de novaĵo de BBC News.

Originala anglalingva teksto	The first global study into the effects of passive smoking has found it causes 600,000 deaths every year. One-third of those killed are children, often exposed to smoke at home, the World Health Organization (WHO) found. The study, in 192 countries, found that passive smoking is particularly dangerous for children, said to be at higher risk of sudden infant death syndrome, pneumonia and asthma. Passive smoking causes heart disease, respiratory illness and lung cancer. "This helps us understand the real toll of tobacco," said Armando Peruga, of the WHO's Tobacco-Free Initiative, who led the study.
Elitrad	La unua tutmonda studo en la efikoj de pasiva fumado trovas, ke estas kaŭzas 600,000 mortojn ĉiun jaron. Unu triono de tiuj mortigita estas infanoj, ofte elmetata al fumi ĉe si, la Monda Organizo pri Sano (KIU) trov. La esploro, en 192 landoj, trovis ke pasiva fumado estas precipe danĝera por infanoj, laŭdire hav pli altan riskon de subita bebo morto sindromo, pneŭmonito kaj astmo. Pasiva fumado kaŭz kormalsano, spira malsano kaj pulmo kancero. "Ĉi tiu helpas nin kompren la vera perdo de tabako," diris Armando Peruga, de la Kiu estas Tabakolibera Iniciato, kiu gvid la esploro.
Trukilo	La unua tutmonda studo en la efikojn de pasiva fumado trovis, ke ĝi kaŭzas 600.000 mortojn ĉiujaran. Unu-tria el tiuj mortigitaj estas infanoj, ofte malkaŝitaj por fumi hejme, la Mondo-sa Organizaĵo (*who) trovis. La studo, en 192 landoj, trovis, ke pasiva fumado estas aparte danĝera por infanoj, diris por esti en pli alta risko de subita infaneto-morta sindromo, *pneumonia *asthma. Pasiva fumado kaŭzas kor-malsanon *respiratoryn malsano *lung kaj kancero. "Ĉi tiu helpoj ni, kiujn komprenas, la vera sonas de *tobacco", diris *Armando *Peruga, de la *who, baaingij *tobacco-Libera Iniciato, kiu kondukas la studon.
Apertium ¹⁷	La unua tutmonda studo en la efikoj de pasiva fumado trovis ĝin kaŭzas 600,000 mortojn ĉiu jaro. Triono de tiuj mortigita estas infanoj, ofte malkovrita fumi ĉe hejmo, la Monda Sana Organiza (WHO) trovita. La studo, en 192 landoj, trovita ke pasiva fumado estas precipe danĝera por infanoj, dirita esti ĉe pli alta risko de subita infana morto sindromo, pneŭmonito kaj astmo. Pasivaj fumadaj aferoj kora malsano, spira malsano kaj pulma kancero. "Tio ĉi helpas nin kompreni la veran perdon de tabako," diris Armando Peruga, de la Kiu estas Tabako-Libera Iniciato, kiu gvidis la studon.
Gramtrans	La unua tutmonda studo en la efikoj de pasiva fumado trovis ke ĝi kaŭzas 600 000 mortojn ĉiun jaron. Unu-triono de tiuj mortigita estas infanoj, ofte eksponitaj por fumi hejme, kiun la Monda Organizaĵo pri Sano (WHO) trovis. La studo, en 192 landoj, trovis ke pasiva fumado estas precipe danĝera por infanoj, diris esti ĉe pli alta risko de enlulila morto, pulminflamo kaj astmo.

¹⁶ Trovita ĉe <<http://gramtrans.com/products/commercial-standard>> la 6an de junio 2011

¹⁷ De <<http://apertium.saluton.dk>>. La ĉefa retejo de Apertium, <<http://apertium.org>>, ne estas ĝisdatigita al sufiĉe nova versio de la motoro, kiun la angla-esperanta paro bezonas

	Pasiva fumado kaŭzas kormalsanon, spiran malsanon kaj pulman kanceron. "Tio helpas al ni kompreni la realan pagon de tabako," Armando Peruga diris, de la Tobacco-Free Initiative de la WHO kiu gvidis la studon.
--	---

Tabelo 9: Diversaj tradukiloj de angla al Esperanto tradukas novaĵartikolon.¹⁸

Kompare kun la aliaj ekzistantaj tradukiloj, kiel Elitrad aŭ Trukilo,¹⁹ ni taksas, ke Apertium havas pli solidan bazon por konstrui pli bonajn tradukilojn. Krome disvolvado super la Apertium-platformo montriĝas multe pli kost-efika el labor-investa vidpunkto. Estas eventuale diskuteble, ĉu ekz. la nuna stato de la angla-esperanta tradukilo de Apertium estas pli aŭ malpli bona ol tiu de Trukilo, sed ne ŝajnas, ke ekzistas io pli-malpli facile solvebla en tio: Apertium estas sufiĉe vasta projekto, verkita kaj ĝisdatigata laŭ la "stato de la arto" de aŭtomata tradukado, kaj donanta multajn eblojn por la diskonigo de la Esperanto-tradukiloj, la ricevado de monhelpoj kaj la eldono de artikoloj en fakrevuoj. Tio estas multe pli malfacile atingebla per unuopulaj klopodoj verkadi Esperanto-tradukilojn.

11. Ebloj kunlabori

Inter Esperanto kaj iu ajn hindeŭropa lingvo kredeble Apertium povas doni kontentigajn rezultojn kontraŭ ne enorma tempoinvesto. Apertium ne estas tre taŭga por traduki inter Esperanto kaj nehindeŭropaj lingvoj (ne estas la celo de Apertium provi okupiĝi pri lingvistike foraj lingvoparoj).

Aparte bone situantaj por krei novajn tradukilojn, pro la akirita sperto, estas la latinidaj lingvoj, kiuj jam havas Apertium-tradukilojn, kiel la portugala, la rumana, la itala, la galega k.a.²⁰ Krome necesas ampleksaj liberaj dulingvaj vortaroj, ekz. kiel okazas inter Esperanto kaj la portugala (danke al la Reta Vortaro kaj Lernu!). Simile eblas por ĝermanaj lingvoj (ĉefe la skandinavaj) kaj slavaj. Principe ŝajnas pli malfacile, pro manko de ampleksaj dulingvaj vortaroj, por keltaj kaj hindiranaj lingvoj.

Sume, ekzistas diversaj formoj kunlabori en la kreado de Esperanto-tradukiloj per Apertium. Certe la plej postula kaj interesa estas evoluigi novan lingvoparon (kiel portugala-Esperanton aŭ ĉeĥa-Esperanton) aŭ novan lingvodirekton (kiel de Esperanto al la hispana). Krome, eblas ankaŭ plibonigi ekzistantan tradukilon, ĉu evoluigante novan version, ĉu atentigante pri ekzistantaj

¹⁸ Novaĵo de <<http://www.bbc.co.uk/news/health-11844169>> la 26an de novembro 2010

¹⁹ Provu ilin en <<http://www.traduku.net/>>.

²⁰ La publikigitajn lingvoparojn de Apertium, en ties nuna stato de evoluo, eblas rigardi ĉe <<http://apertium.svn.sourceforge.net/viewvc/apertium/trunk/>>; la nepublikigitajn eblas rigardi ĉe <<http://apertium.svn.sourceforge.net/viewvc/apertium/staging/>> (paroj en progresinta stadio de evoluigado), <<http://apertium.svn.sourceforge.net/viewvc/apertium/nursery/>> (paroj en komenca stadio de evoluigado) kaj <<http://apertium.svn.sourceforge.net/viewvc/apertium/incubator/>> (ĉefe eroj por evoluigi lingvoparojn, kiel vortaroj).

eraroj.²¹ Kaj certe, libera disponigo de dulingvaj vortaroj estas ege dankinda: ja evoluigantoj bezonas kiel eble plej ampleksajn kaj fidindajn vortarojn, terminarojn, esprimarojn ktp.

Bibliografio

ALÒS i FONT, Hèctor (red.). *Diccionari Apertium català-esperanto*. Shupashkar, 2010.

<<http://www.esperanto.cat/vortaro/apertium-ca-eo.pdf>>

FORCADA, Mikel L.; BONEV, Boyan Ivanov & ORTIZ ROJAS, Sergio. *Documentation of the Open-Source Shallow-Transfer Machine Translation Platform Apertium*. [Alacant:] Departament de Llenguatges i Sistemes Informàtics, 2008.

<<http://xixona.dlsi.ua.es/~fran/apertium2-documentation.pdf>>

FORCADA, Mikel L.; GINESTÍ-ROSELL, Mireia & NORDFALK, Jacob. Apertium: a free/open-source platform for rule-based machine translation. *Machine translation*, aperonta.

WIECHETEK, Linda; TYERS, Francis M. & OMMA, Thomas. Shooting at flies in the dark: Rule-based lexical selection for a minority language pair. *Lecture Notes in Artificial Intelligence* vol. 6233, p. 418-429, 2010. <<http://xixona.dlsi.ua.es/~fran/publications/icetal2010.pdf>>

²¹ Por atentigi pri eraroj, legu la instrukciojn ĉe <http://wiki.apertium.org/wiki/Esperanto/Outstanding_tests>.