

BI MAILATAKO MORFOLOGIAREN EUSKARARAKO EGOKITZAPENA

A. Agirre, E. Agirre, I. Alegria, X. Arregi, X. Artola,
A. Diaz de Ilarraza, P. Goenaga, M. Maritxalar, K. Sarasola
Informatika Fakultatea - E.H.U.- U.P.V. - DONOSTIA
M. Urkia

U.Z.E.I. - Unibertsitate-Zerbitzueterako Euskal Ikastetxea - DONOSTIA.

TOPIC AREA: Morphological Analysis

ABSTRACT

The simplicity of English inflections made for reduced interest in research on morphological analysis by computer. In English, the most common practice is to use a lexicon of all of the inflected forms or a minimum set of morphological rules. Which means that while a great many language independent tools have been developed for syntactic and semantic analysis, the same cannot be said for morphological tools.

In 1981, Kaplan and Kay made a valuable contribution in designing a formalism for phonological generation by means of rules compiled in an automaton. This idea would later be followed up by Koskenniemi, who in 1983, designed the computational model for two-level morphology.

This paper describes the formalism of the two-level morphology and its application to Basque. The morphological analyzer presented has been designed with the aim of laying the foundations for further development of automatic processing of Basque. In this context, the morphological analyzer is used in the elaboration of the XUXEN spelling checker/corrector.

BI MAILATAKO MORFOLOGIAREN EUSKARARAKO EGOKITZAPENA

A. Agirre, E. Agirre, I. Alegria, X. Arregi, X. Artola,
A. Diaz de Ilarraza, P. Goenaga, M. Maritxalar, K. Sarasola
Informatika Fakultatea - E.H.U.- U.P.V. - DONOSTIA
M. Urkia

U.Z.E.I. - Unibertsitate-Zerbitzuetarako Euskal Ikastetxea - DONOSTIA.

ARLOA : Analisi Morfologikoa

LABURPENA

Analisi morfologiko automatikorako proposaturiko formalismoen artean bi mailatakoa izan da azken urteotan interes gehien piztu duena. Duen ezaugarririk behinena mailen arteko bereizketa da, azalekoa edo testukoa bata eta lexikoa edo hiztegikoa bestea. Bi maila horien arteko aldaketak erregela bidez adierazten dira, inplementatzerakoan egoera finituzko automata bilakatuko diren erregelez, hain zuzen ere. Programa nagusia erabiltzen den hizkuntzarekiko independentea da eta baliagarria analisi nahiz sorkuntzarako.

Artikulu honetan formalismo hau eta bere euskararako egokitzapena deskribatzen da, tresna linguistikoak sortzeari zuzendutako proiektu baten barnean. Testuinguru honetan, analizatzaile morfologikoaren azpiproduktu gisa, euskararako zuzentzaile ortografiko baten garapenean dihardugu (*).

(*) Honako lan hau Gipuzkoako Foru-Diputazioak babestutako proiektu batean kokatu behar da.

BI MAILATAKO MORFOLOGIAREN EUSKARARAKO EGOKITZAPENA

A. Agirre, E. Agirre, I. Alegria, X. Arregi, X. Artola,
A. Diaz de Ilarraza, P. Goenaga, M. Maritxalar, K. Sarasola
Informatika Fakultatea - E.H.U.- U.P.V. - DONOSTIA
M. Urkia

U.Z.E.I. - Unibertsitate-Zerbitzueterako Euskal Ikastetxea - DONOSTIA.

1. SARRERA

Ingelesak flexiorako duen sinpletasunagatik edo, ordenadorezko analisi morfologikoak ez du aparteko interesik piztu ikertzaileen artean. Ingelesez, forma flexionatu guztiak biltzen dituzten lexikoak erabiltzen dira normalean, edo erregela morfologiko gutxi batzuk bestela [Winograd 83]. Honen kariaz, hizkuntzarekiko independenteak diren analisi sintaktiko nahiz semantikorako tresna linguistikoak ugalduta diren bitartean, tresna morfologikoek ez dute apenas atentziorik jaso.

Nolanahi ere, badira ingelesa ez bezalako hizkuntzentzat garatu diren ordenadorezko analisi morfologikorako sistemak, hala nola, GETAko ATEF [GETA 82], suomierarako sistemak [Broda et al. 80] eta beste batzuk.

1981ean Kaplan eta Kay-ek [Kaplan et al. 81] sorkuntza fonologikorako formalismoa diseinatu zuten. Formalismo honek automatetara konpilatzen diren erregelak darabiltza eta, ideia interesgarri honi jarraikiz, Koskenniemi-k hemen aztertuko dugun bi mailatako formalismoa diseinatu zuen.

Azken urteotan analisi morfologikorako beste zenbait formalismo ere garatu da. Gaztelaniarako, esate batera, MARS [Meya 87], deskonposaketa

morfologikoan egoera finituzko automata erabiltzen duena, eta AM [Martí 87], baldintzez osatutako automata Markoviarraz baliatzen dena.

Artikulu honetan bi mailatako morfologiaren formalismoa eta euskararentzat egindako egokitzapena deskribatzen dira.

2. BI MAILATAKO MORFOLOGIA

1983an Koskenniemi-k [Koskenniemi 83] bi mailatako morfologiaren eredu konputazionala definitu zuen. Eredu honek harrera bikaina jaso du ondorengo urteetan, besteak beste, dituen ezaugarri hauengatik:

- Eredu orokorra da, edozein hizkuntzari aplikatu dakiokiena.
- Baliagarria da hitzen analisi morfologikorako zein hitz-sorkuntzarako.
- Ezagutza linguistikoa eta algoritmoa bereizi egiten ditu eta, ondorioz, programa berak edozein hizkuntzatarako balio dezake.
- Analizatu edo sortuko den hitzaren azaleko maila eta hiztegiko sisteman (sistema lexikoan) errepresentatzen den maila lexiko edo sakonekoa argi eta garbi bereizten ditu. Hau dela eta, ez dago aldaketa morfofonologikoengatik sortutako morfema baten forma desberdinak gorde beharrik.
- Fonologia sortzaileko berridazketa-erregelen ordean erregela paraleloak erabiltzen ditu, sistema kontzeptual zein konputazionalki errazago bihurtuz.

Morfologia honek oinarrizko bi osagai ditu: erregelak eta sistema lexikoa.

2.1. ERREGELAK

Bi mailatako ereduak errepresentazio lexikoa eta azalekoa erabiltzen ditu. Lexikoak erroen eta afixuen errepresentazio morfofonologikoak dauzka.

Bi errepresentazioen artean ez dago tarteko egoerarik, eta haxe da fonologia sortzailearekiko diferentzia nagusia. Beraz, hitzen azterketa azaleko formari dagozkion errepresentazio lexiko onargarriak aurkitzean datza. Alderantziz gertatzen da sorkuntzan, errepresentazio lexiko ezagunetik abiatu eta berari dagozkion azaleko errepresentazioak bilatzen bait dira.

Erregelek hiru zati dituzte:

- Korrespondentzia, edo karaktere-bikote bat, non lehenengoa lexiko-mailakoa bait da eta bigarrena azaleko mailan aurrekoari dagokiona. Karaktere hauek konketu nahiz abstraktuak izan daitezke, azken hauek erregelen jeneralizazioa ahalbidetzen dutelarik. Karaktere abstraktuen adibide gisa aipatu litezke C edozein kontsonante adierazteko, V edozein bokal, e.a.
- Testuingurua, korrespondentzia gertatzen deneko kasuak mugatzen dituena, aurreko eta ondorengo karaktereen arabera.
- Eragilea, testuinguruaren eta korrespondentzian adierazitako bikotearen artean zer-nolako erlazioa dagoen finkatzen duena. Era desberdinetakoa izan daiteke: testuinguru-murriztapena (\Rightarrow), azaleko koertzioa (\Leftarrow), biak batera (\Leftrightarrow) edo debeku-ezarpina (\nrightarrow).

Erregela hauen hasierako sintaxiak aldaketa batzuk izan ditu ([Koskeniemi 85], [Dalrymple et al. 87]) konpiladore bat inplementatu ahal izateko eta, honela, automatetarako itzulpena eskuz egin behar ez izateko.

Adibidez, "l:i \Rightarrow b:b _ e:e" erregelak, "l:i" bikoteak (hau da, "l" karakterearen agerpena maila lexikoan eta "i" karakterea azalekoan) ezkerreko

testuinguruan "b:b" bikotea eta eskuinekoan "e:e" bikotea eskatzen duela adierazten du. "l:i \Leftarrow b:b _ e:e" erregelak, berriz, "b:b" ezkerrean, "e:e" eskuinean eta "l" karakterea maila lexikoan ageri den edonon azaleko karaktereak derrigorrez "i" izan behar duela adierazten du.

l:i	\Rightarrow	b:b	_	e:e
korrespondentzia	eragilea	ezkerreko		eskuineko
		testuingurua		testuingurua

Mako eta barren erabilpenak testuinguru alternatibo desberdinak definitzeko aukera eskaintzen du korrespondentzia baterako. Adibidez

k:g \Leftrightarrow [l&2:l00 / n&2:n00 / n4\$:n00 / n\$:n0] _ o:o

ondoko lau erregelen baliokidea da:

k:g \Leftrightarrow l&2:l00 _ o:o

k:g \Leftrightarrow n&2:n00 _ o:o

k:g \Leftrightarrow n4\$:n00 _ o:o

k:g \Leftrightarrow n\$:n0 _ o:o

Erregelak erabili ahal izateko premiazkoa da elementu hauek definitzea:

- Alfabetoa, edo azaleko karaktereen multzoa. Øak karaktere hutsa denotatzen du.
- Erregeletan erabiltzen diren azaleko alfabetoaren azpimultzoak.
Esate baterako, Ck edozein kontsonante denotatzen du eta Vk edozein bokal.
- Lexikoko alfabetoa. Erregelen aplikazioa kontrolatzen duten hautapen-marka (\$, §, 4, 5,...) izeneko karaktere bereziek eta azaleko alfabetoak osatzen dute.

- Lexikoko alfabetoaren azpimultzoak. Azalekoen parekoak dira.
- Laburduren edo erregelaren azpiadierazpenen definizioak.
- Erregelak berak.

2.2. SISTEMA LEXIKOA

Erregelek bi mailen arteko desberdintasunak deskribatzeko balio duten bitartean, sistema lexikoak morfema-multzoa definitzen du, morfemen artean egon daitezkeen kateamenduen arabera sailkapena eginez. Azpilexikoen multzoa eta erroen eta afixuen sekuentzia posibleak erregulatzen dituzten jarraitze-klaseek sistema hau osatzen dute.

Azpilexikoek ezaugarri berdineko elementu lexikoak (atzizkiak, aurrizkiak, izen-lemak, aditz-lemak...) biltzeko balio dute. Azpilexiko guztiek egitura bera dute, identifikatzen dituen izena eta sarrera-multzoa. Sarrera bakoitzak hiru eremu ditu :

- Adierazpen lexikoa, alfabeto lexikoko karaktere-sekuentzia bat da. Karaktere hauek azaleko karaktereak edo hautapen-markak izan daitezke. Azkeneko hauei azaleko beste karaktere batzuk egokitu lekizkieke erregelaren bitartez.
- Dagokion jarraitze-klasea. Zenbait azpilexiko edota beste jarraitze-klase batzuk biltzen dituen identifikadorea da. Jarraitze-klasean biltzen diren osagaiak dira definitutako sarreraren atzetik ager daitezkeen bakarrak.
- Sarrerari dagokion informazio morfologikoa.

Beraz, jarraitze-klaseak hitz batean ager daitezkeen morfemen arteko konbinaketa posibleak definitzeko mekanismoaren oinarria dira. Beraien adierazte-ahalmena txikia da eta, honengatik, kasu batzutan beharrezkoa izango

ez litzatekeen zenbait deskripzio-bikoizketa gertatu ohi da, morfemen arteko distantzia handiko menpekotasuna tratatzerakoan, esaterako.

2.3. PROGRAMA

Programak bi modulu laguntzaile nagusi erabiltzen ditu, Fsp eta Lex. Lehenengoak automataren lana egiten du, sarrera gisa karaktere-bikoteak onartzen dituelarik. Horretarako, hasieraketan erregela bakoitzari dagozkion automatetatik egoera finituzko automata bakarra osatzen du, lerrokatu eta onargarri diren bikoteen multzoa lortzen du.

Lex modulu lexikoak lexiko-atzipenaren funtzioa egiten du. Honetarako, arbola moduan antolatutako azpilexikoetan banaturiko fitxategi bat dago sortuta, arbolaren adar bakoitza karaktere bat delarik, atzipen inkrementala lortuz. Azpilexikoen arteko lotura jarraitze-klaseen bidez definitzen da.

Modulu hauek hizkuntzarekiko independenteak dira eta analisisa nahiz sintesia egin dezakeen programa nagusi batetik deituak dira.

Ereduaren konplexutasun konputazionala sakonki aztertzen da [Karttunen 83] eta [Barton 85]-en. Biak bat datoz hizkuntzaren konplexutasunak analisi edo sorkuntzaren abiaduran eragin adierazgarririk ez duela diotenean.

3. EUSKARARAKO APLIKAZIOA

3.1. EUSKARAREN MORFOLOGIAREN DESKRIBAPEN LABURRA

Euskara hizkuntza eranskaria da, hau da, hitzak osatzeko hiztegi-sarrerak independenteki hartzen ditu funtzio desberdinetarako (kasua barne) beharrezkoak diren elementuak. Konkretuki determinatzaileari, numeroari eta

deklinabide-kasuari dagozkien afixuak ordena honetan eta independenteki hartzen dira.

Euskararen ezaugarri nagusienetako bat, inguruko hizkuntzetatik bereizten duena, kasuak dituen deklinabide-sistema baten bidez funtzionatzea da. Determinatzailea, numeroa eta kasuaren flexioak izen-sintagma osoari aplikatzen zaizkio eta ez sintagmaren osagai bakoitzari, inguruan ditugun hizkuntza erromanikoetan gertatzen den bezala. Euskal deklinabidea bakarra da, hau da, sarrera guztiei eransten zaien deklinabide-taula bakarra dauka.

Euskaraz aditz-sistema, hitz-elkarketa eta eratorpena aberatsak dira. Horrek sarrera bakar batetik abiatuz sorkuntza lexikoa ahalbideratzen du.

3.2. ZERGATIK BI MAILATAKO MORFOLOGIA ?

Euskararen morfologia aberatsa denez formalismo ahaltzu bat aukeratu behar zen bere automatizazioari ekiteko. Hasieran azaldu ditugu "bi mailatako morfologia" formalismoaren ezaugarriak, baita berridazketa-erregelekiko dauzkan abantailak ere. Gainera, azken urteotan bi mailatako morfologia arrakastaz aplikatu zaio hainbat hizkuntzari, morfologikoki hain korapilatsu diren suomiera edo arabiera bezalako hizkuntzak ere tartean. Suomieraz, hizkuntza eranskaria, euskaraz gertatzen denaren antzekoa aurkitzen dugu. Bi hizkuntzek tratamendu antzekoa jaso dezakete. Dena den, euskararen kasuan unitate lexiko bakoitzak erro bakarra duenez sinpleagoa da; suomieraz, ordea, hau ez da beti gertatzen. Deklinabideak antzeko funtzionamendua dauka bi hizkuntzetan, baita hitz-elkarketak eta eratorpenak ere.

Lexiko-maila : sakon **2ko**

Azaleko maila : sakon **eko**

1. erregela. "r" epentetikoa erantsi.

1:r ⇔ V:V _ (erregela sinplifikatua)

Besteak beste, lexiko-mailan 1 hautapen-marka (1ak "r" epentetikoa eskatzen duten atzizkiak bereizten ditu) bokal baten ondoren badago, lexikoko 1ari azaleko "r" a dagokiola adierazten du. Gainera, inplikazioa bikoitza denez, **1:r** bikotea ageri den aldioro aurreko karakterea bokal bat izango da bi mailetan.

Adibidea: **semeri**

Lexiko-maila : seme **1i**

Azaleko maila: seme **ri**

b) Ezabaketa : Bi elementu elkartzen direnean karaktere baten desagerketa:

Kasu honetan, lekuzko izen propio batzuk destinatiboan jasaten duten "a" artikularen galera deskribatzen da.

4. erregela. "a"ren galera.

a:Ø ⇔ _ #\$:ØØ [a:a / e:e / o:o] ;:Ø [2ra:Øra / 2t:Øt / 2ko:Øko / \$ko:Øko] #2a:ØØa
--

Erregela honek dioenez, lexiko-mailan "a" karakterearen segidan ";2ra" badator, azaleko mailan "a" artikuluari karaktere hutsa dagokio (Ø) eta ";2ra"ri "ØØra" katea (";" hautapen-marka "a" artikulua galtzen duten lekuzko izen propioek daramate eta 2 marka, berriz, "e" epentetikoa eskatzen dutenek, baina azken hau ez da hemen aplikatzen, aurrekoa bokala zen eta).

Adibidea: **Azpeitira**

Lexiko-maila: Azpeitia; 2ra

Azaleko maila: AzpeitiØØ Øra

6. erregela. Bi kontsonante leherkari gor elkarren ondoan suertatzean, lehenengoaren desagerketa (deklinabide-kasuak elkartzean).

$k:\emptyset \Leftrightarrow _ [\$k:\emptyset k / \$t:\emptyset t]$

etxetiko

Lexiko-maila: etxetik \$ko

Azaleko maila: etxetiØØko

c) Aldaketa : Kontsonante gorra ozen bilakatzea sudurkari baten eraginez, etab.; bokalen arteko aldaketak.

7. erregela. "k" gorra "g" ozen bihurtzen da.

$k:g \Rightarrow [l\&2:l\emptyset\emptyset / n\&2:n\emptyset\emptyset / n4\$:n\emptyset\emptyset / n\$:n\emptyset] _ o:o$

Usurbilgo

L.M.: Usurbil\$ 2ko

A.M.: Usurbil ØØgo

5. erregela. "a" bokala "e" bihurtzea.

$a:e \Leftrightarrow _ 7\$: \emptyset\emptyset [a:a / e:e / o:o]$
 $\%2:\emptyset\emptyset [n:n / l:l]$
 $72:\emptyset\emptyset [k:k / tik:tik / an:an / z:z]$

den

L.M.: da% 2n

A.M.: deØ Øn

aterea

L.M.: atera7 \$a

A.M.: atereØ Øa

3.3.2. Sistema lexikoa

Errepresentazio lexikoa definitzeko, sarrera bakoitzari bere azpilexikoa eta jarraitze-klasea egokitzen zaio.

a) Azpilexikoak: Hiru multzotan banatu ditugu: lemak (hiztegiko sarrerak), aditz laguntzaileak eta "egitura morfotaktiko" deitutakoak. Azken hauek aurrizkiz eta atzizkiz osatzen dira, deklinabide-kasuak direla, plurala, aditz-atzizkiak, etab.

Sarrera guzti hauen multzoa azpilexikotan banatuta adierazten da. Azpilexikoetako sarrera guztiek beren jarraitze-klasea eta informazio morfologikoa daramate. Informazio morfologiko gisa, kategoria, deklinabide-kasua, numeroa, denbora-modua, aditz-erroa, aditz-mota, pertsona gramatikala, eta sarrera bakoitzari erants dakioken informazio espezifikoak gordetzen da.

Zenbait kasutan informazio morfosintaktikoa ere gehitu da, aurrerago ikusi ahal izango den bezala (3.3.2.c).

b) Jarraitze-klaseak: Ezinezkoa da beti jeneralizatzea. Esate baterako, izen eta adjektiboak jarraitze-klase banatan bildu badira ere, beste zenbait kategoriatako elementuek, adberbioek edo aditzek kasu, soluzio berezituagoa behar izan dute. Adib.: "polit" sarrerari ADJ jarraitze-klasea dagokio. Honela definituko dugu ADJ klasea:

$$\text{ADJ} = (\text{I1}, \text{ago_egi_en})$$
$$\text{I1} = (\text{a_ms}, \text{IMG}, \text{IMS}, \text{IMP})$$

non, maiuskulaz dauden identifikadoreak jarraitze-klaseak diren eta minuskulazkoak, aldiz, azpilexikoak.

Analisi morfologikoa, horretara, elementuen deskonposaketa burutu ahala hiztegitik datuak atera eta informazio hori biltzean datza.

c) Informazio morfosintaktikoa: Koskenniemi-ren eredua morfologiara mugatua badago ere, euskararako aplikazio honetan syntaxira hurbiltzeko saio bat egin da. Oraingoan sistematizaziorik ez dago, baina menpekotasun-erlazioa,

sortarazten duen kategori aldaketa eta beste zenbait puntu aztertu dira.

Aipatu erlazioa, adibidez, menderagailuek markatzen dute:

du3 LAT

LAT= (IØ, n1, la)

non n1(ERL ERLT)

la (ERL KNPL) diren.

Berehalako helburua da testu-hitz bateko osagai bakoitzaren informazioa jaso eta multzo osoarena lortzea, baina oraindik ezin dira ondorio zehatzak aurkeztu. Berdin kategoria funtzionalei dagokienez ere, sintagma bateko azken osagaiak emaitza osoa eta bateratua emango bait du.

Kategori aldaketak jarraitze-klaseetako osagaiei esker gauzatu ahal izan dira. Honela, aditz bat izen bilaka daiteke (egin-egite) edo izen eta adjektiboak aditz (gizon-gizondu, txiki-txikitu).

3.4. SISTEMAREN ABANTAILAK IKUSPEGI LINGUISTIKOTIK

Koskenniemi-ren eredua lexiko-mailan oinarritzen da eta bi mailen arteko korrespondentzia definitzeko erregelez baliatzen da. Beraz, azpilexikoetan formak oso-osorik, aldaketarik gabe, gordetzen dira, beste sistemetan ez bezala.

Adibidez, "ama" unitate lexikoa da, "a" organikoa du, baina datibo plurala "amei" forma da; eta horregatik sistema askotan "ama" erabili ordez "am" erabili beharko litzateke. Bi mailatako morfologian, lexiko eta azaleko maila argi eta garbi bereizten direnez, unitatea oso-osorik eta alferrikako bikoizketarik gabe mantentzeko dezakegu. Linguistikaren ikuspuntutik garrantzi handikoa deritzogu analisi morfologikoaren planteamendu honek dakarren unitate lexikoarekiko errespetoari.

4. ONDORIOAK

Sistema hau lehenengoz suomierarako eta geroztik beste zenbait hizkuntzatarako erabili da. Suomierarako inplementazioaz gain, ingelesarako [Karttunen et al. 83] eta arabierarako estensiboki inplementatuta dago, eta frantseserako [Lun 83], japonierarako [Alam 83] eta errumanierarako [Khan 83] prototipoak landu dira. Linguistei edozein hizkuntzaren morfologia askatasun osoz adierazteko aukera eskaintzen die, hizkuntza jakinetan moldatutako programa zurrunei eta aspektu operatiboei itzuri eginez.

Burroughs B7800 [Koskenniemi 83] batean eginiko lehen inplementazioko programak Pascaleko 2000 lerro zeuzkan eta analisisan, batazbeste, 0,1 sg. behar zituen formako. Gure proiektuan C lengoia inplementatu da, PC eta bateragarrietan erabili ahal izateko, baina ez ditugu oraindik programaren eraginkortasunari buruzko emaitzak.

Lehenxeago aipatu dugunez, analizatzaile hau euskararako zuzentzaile ortografiko batean erabiliko da lehendabizi: flexio-maila handiko hizkuntzek arazoak izaten dituzte formak hiztegian gordetzerakoan eta arazo hauek konpontzeko premiazkoa da hitzak morfologikoki behar bezala lantzea. Zuzentzaileak ontzat emango du deskonposaketa morfologiko zuzena onartzen duen edozein hitz. Deskonposaketa posible guztiak lortzea izango da analizatzailearen eginkizuna.

Analizatzaile-sortzaile morfologikoa etorkizunerako oinarritzko tresnatzat hartzen dugu, lengoia naturalaren tratamenduan burutu asmo ditugun lanetarako behar-beharrezkotzat.

ERREFERENTZIAK:

- [Agirre et al. 89] Agirre E., Alegria I., Arregi X., Artola X., Diaz de Ilarraza A., Sarasola K., Urkia M. *Aplicación de la morfología de dos niveles al euskara*. SEPLN 8. zb., 87-102 or., 1989.
- [Alam 83] Alam, Y. S. . *A two-level analysis of Japanese*. Texas Linguistic Forum, vol. 22, Pp. 229-252, 1983.
- [Barton 85] Barton, E. . *Computational Complexity in two-level Morphology*, 1985.
- [Broda et al. 80] Brodda, B. and F. Karlsson. *An experiment with Automatic Morphological Analysis of Finnish*. Papers from the Institute of Linguistics, University of Stockholm. Publication 40, 1980.
- [Dalrymple et al.87] Dalrymple M., Kaplan R.M., Karttunen L., Koskenniemi K., Shaio S., Wescoat M. *Tools for Morphological Analysis*. Report N. CSLI-87-108, 1987.
- [Euskaltzaindia 73] Euskaltzaindia. *Aditz laguntzaile batua*. Euskaltzaindia, Bilbo 1973.
- [Euskaltzaindia 85] Euskaltzaindia. *Euskal Gramatika: Lehen urratsak (I eta II)*. Euskaltzaindia, Bilbo 1985.
- [GETA 82] G.E.T.A. . *Le point sur ARIANE-78 debut 1982*. Vol.1, partie 1: Le logiciel. pp. 28-75, 1982.
- [Kaplan et al. 81] Kaplan, Ronald M., and M. Kay. *Phonological rules and finite-state transducers*. Paper read at the annual meeting of the Linguistic Society of America in New York City, 1981.
- [Karttunen et al. 83] Karttunen, L., and K.Wittenburg. *A two-level analysis of English*. Texas Linguistic Forum, vol. 22, Pp. 217-228,1983.
- [Karttunen 83] Karttunen, L. . *KIMMO : A two-level Morphological Analyzer*. Texas Linguistic Forum, vol. 22, Pp.165-186, 1983.

- [Kay 73] Kay, Martin. *Morphological Analysis*. A.Zampolli & N. Calzolari eds. (1980). Proc. of the Int. Conference on Computational Linguistics (Pisa), 1973.
- [Khan 83] Khan, Robert. *A two-level analysis of Rumanian*. Texas Linguistic Forum, vol. 22, Pp. 253-270, 1983.
- [Koskenniemi 83] Koskenniemi, K. . Two-level Morphology: A general Computational Model for Word-Form Recognition and Production, University of Helsinki, Department of General Linguistics. Publications 11, 1983.
- [Koskenniemi 85] Koskenniemi, K. *Compilation of Automata from Morphological Two-level Rules*. Pp. 143-149. Publication 15. University of Helsinki, 1985.
- [Lun 83] Lun, S. *A two-level analysis of French*. Texas Linguistic Forum, vol. 22, Pp. 271-278, 1983.
- [Martí 87] M. A. Martí. *Un sistema de análisis morfológico por ordenador*. SEPLN 4. zb., 1987.
- [Meya 87] M. Meya. *Análisis morfológico como ayuda a la recuperación de información*. SEPLN 4. zb., 1987.
- [Ritchie et al. 87] Ritchie, G.D., S.G. Pulman, A.W.Black and G.J. Russell. A Computational Framework for Lexical Description. Computational Linguistics, vol. 13, numbers 3-4, 1987.
- [Sarasola 82] Sarasola, Ibon. *Gaurko euskara idatziaren maiztasun-hiztegia*. (3gn. liburukia), GAK, Donostia, 1982.
- [Winograd 83] Winograd, Terry. *Language as a cognitive process*. Vol. 1: Syntax, pp 544-549. Addison-Wesley, 1983.