# D2.3.mt.mt
# Language Report for Maltese
# (Maltese version)

Version 1.0

2011-08-03

METANET4U

www.metanet4u.eu

The central objective of the Metanet4u project is to contribute to the establishment of a pan-European digital platform that makes available language resources and services, encompassing both datasets and software tools, for speech and language processing, and supports a new generation of exchange facilities for them.
This central objective is articulated in terms of the following main goals:
**Assessment**: to collect, organize and disseminate information that permits an updated insight into the current status and the potential of language related activities, for each of the national and/or language communities represented in the project. This includes organizing and providing a description of: language usage and its economic dimensions; language technologies and resources, products and services; main actors in different areas, including research, industry, government and society in general; public policies and programs; prevailing standards and practices; current level of development, main drivers and roadblocks; etc.
**Collection**: to assemble and prepare language resources for distribution. This includes collecting languages resources; documenting these language resources; upgrading them to agreed standards and guidelines; linking and cross-lingual aligning them where appropriate.
**Distribution**: to distribute the assembled language resources through exchange facilities that can be used by language researchers, developers and professionals. This includes collaboration with other projects and, where useful, with other relevant multi-national forums or activities. It also includes helping to build and operate broad inter-connected repositories and exchange facilities.
**Dissemination**: to mobilize national and regional actors, public bodies and funding agencies by raising awareness with respect to the activities and results of the project, in particular, and of the whole area of language resources and technology, in general.

METANET4U is a project in the META-NET Network of Excellence, a cluster of projects aiming at fostering the mission of META. META is the Multilingual Europe Technology Alliance, dedicated to building the technological foundations of a multilingual European information society.

META

METANET4U is co-funded by the participating institutions and the ICT Policy Support Programme of the European Commission

and by the participating institutions:

Faculty of Sciences, University of Lisbon

Instituto Superior Técnico

University of Manchester

University *Alexandru Ioan Cuza*

Research Institute for Artificial Intelligence, Romanian Academy

University of Malta

Technical University of Catalonia

Universitat Pompeu Fabra

Revision History

| Version | Date | Author | Organisation | Description |
|---|---|---|---|---|
| 1.0 | 2011-08-03 | Mike Rosner, Jan Joachimsen | UOM | Final version |

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

# D2.3.mt.mt Language Report for Maltese
## (Maltese version)

Document METANET4U-2011-D2.3.mt.mt
EC CIP project #270893

Deliverable
Number: D2.3.mt.mt
Completion: Final
Status: Submitted
Dissemination level: Public

Responsible: Asunción Moreno (WP2 coordinator)

Contributing Partners: UOM

Authors: Mike Rosner, Jan Joachimsen

Translator: Ritianne Stanyer

Reviewer: Bernard Cachia

© all rights reserved by FCUL on behalf of METANTE4U

# Tabella tal-kontenut

**Sommarju Eżekuttiv ..................................................................... 3**

**Riskju għal-Lingwi Tagħna u Sfida għat-Teknoloġija Lingwistika ..................................................................... 4**

Il-Konfini Lingwistiċi jfixklu s-Soċjetà Ewropea tal-Informazzjoni ..................................................................... 5

Il-Lingwi Tagħna f'Riskju ..................................................................... 5

It-Teknoloġija Lingwistika hija Teknoloġija Katalizzanti Ewlenija ..................................................................... 6

Opportunitajiet għat-Teknoloġija Lingwistika ..................................................................... 7

Sfidi li t-Teknoloġija Lingwistika Taffaċċja ..................................................................... 8

Il-Ksib tal-Lingwi ..................................................................... 8

**Il-Malti fis-Soċjetà tal-Informazzjoni Ewropea ..................................................................... 10**

Fatti Ġenerali ..................................................................... 10

Partikolaritajiet tal-Lingwa Maltija ..................................................................... 11

Żviluppi riċenti ..................................................................... 13

Il-Kultivazzjoni tal-Lingwa f'Malta ..................................................................... 14

Il-Lingwi fl-Edukazzjoni ..................................................................... 15

Aspetti internazzjonali ..................................................................... 18

Il-Malti fuq l-Internet ..................................................................... 18

Referenzi ..................................................................... 22

**Appoġġ ta' Teknoloġija Lingwistika għall-Malti ..................................................................... 24**

Teknoloġiji Lingwistiċi ..................................................................... 24

Arkitetturi ta' Applikazzjonijiet għat-Teknoloġija Lingwistika ..................................................................... 24

Oqsma ewlenin ta' applikazzjoni ..................................................................... 25

It-Teknoloġija Lingwistika 'wara l-kwinti' ..................................................................... 32

It-Teknoloġija Lingwistika fl-Edukazzjoni ..................................................................... 34

Programmi ta' Teknoloġija Lingwistika ..................................................................... 36

Tabella ta' Għodod u Riżorsi għall-Malti ..................................................................... 39

Konklużjonijiet ..................................................................... 40

Referenzi ..................................................................... 42

**Dwar META-NET ..................................................................... 45**

Linji ta' azzjoni ..................................................................... 46

Organizzazzjonijiet Membri ..................................................................... 47

Referenzi......................................................................................................................................50

## Sommarju Eżekuttiv

Bosta lingwi Ewropej qegħdin f'riskju, li jispiċċaw vittmi tal-era diġitali minħabba li mhumiex rappreżentati biżżejjed u għandhom nuqqas ta' riżorsi fuq l-internet. Illum ħafna opportunitajiet fis-suq reġjonali jibqgħu ma jiġux sfruttati minħabba ostakli lingwistiċi. Jekk ma niħdux azzjoni issa, ħafna ċittadini Ewropej se jsiru soċjalment u ekonomikament żvantaġġjati għax jitkellmu l-lingwa nattiva tagħhom.

Teknoloġija lingwistika innovattiva (TL) hija intermedjarju li se jippermetti liċ-ċittadini Ewropej jipparteċipaw f'soċjetà tal-informazzjoni ugwali u inklussiva u b'għarfien li jwassal għal suċċess ekonomiku. It-teknoloġija lingwistika multilingwi se tkun portal għal komunikazzjoni instantanja, irħisa u mingħajr sforzi u mezz ta' interazzjoni bejn konfini lingwistiċi.

Illum, is-servizzi lingwistiċi huma offruti primarjament minn fornituri kummerċjali mill-Istati Uniti. Google Translate, servizz mingħajr ħlas, huwa eżempju wieħed biss. Is-suċċess riċenti ta' Watson, sistema tal-kompjuter IBM li rebħet l-episodju tal-programm tal-logħob *Jeopardy* kontra kandidati umani, turi l-potenzjal enormi tat-teknoloġija lingwistika. Bħala Ewropej, għandna nsaqsu lilna nfusna bosta mistoqsijiet urġenti:

- Il-komunikazzjoni u l-infrastruttura tal-għarfien tagħna għandhom ikunu dipendenti fuq kumpaniji monopolistiċi?
- Nistgħu verament niddependu fuq servizzi marbuta mal-lingwa li jistgħu jintfew immedjatament minn oħrajn?
- Qegħdin nikkompetu b'mod attiv fis-suq globali għal riċerka u żvilupp fit-teknoloġija lingwistika?
- Partijiet terzi minn kontinenti oħra huma lesti biex jindirizzaw il-problemi tagħna ta' traduzzjoni u kwistjonijiet oħra marbutin mal-multilingwiżmu Ewropew?
- L-isfond kulturali Ewropew tagħna jista' jgħin biex jibni s-soċjetà ta' għarfien billi joffri teknoloġija aħjar, aktar sikura, aktar preċiża, aktar innovattiva u aktar robusta u ta' kwalità għolja?

Din il-*White Paper* għal-lingwa Maltija turi li f'Malta hemm potenzjal għal industrija ta' teknoloġa lingwistika u ambjent ta' riċerka qawwija. Minkejja li jeżistu għadd ta' teknoloġiji u riżorsi għall-Malti, hemm bil-wisq anqas teknoloġiji u riżorsi għal-lingwa Maltija milli ghal lingwi Ewropej "akbar" u ċertament mhux biżżejjed biex jissapportjaw il-firxa sħiħa ta' applikazzjonijiet sensittivi għal-lingwi li huma disponibbli għal dawk il-lingwi oħra.

Skont l-assessjar iddettaljat f'dan ir-rapport għandha sseħħ azzjoni immedjata qabel ikunu jistgħu jinkisbu xi riżultati importanti għal-lingwa Maltija.

## Riskju għal-Lingwi Tagħna u Sfida għat-Teknoloġija Lingwistika

*Bħalissa qed naraw rivoluzzjoni diġitali li tista' titqabbel mal-invenzjoni tal-istampar ta' Gutenberg.*

Qed naraw rivoluzzjoni diġitali li qed tħalli impatt b'mod drammatiku fuq il-komunikazzjoni u s-soċjetà. Żviluppi riċenti fit-teknoloġija tal-komunikazzjoni diġitizzata u tan-networks xi drabi jiġu mqabbla mal-invenzjoni tal-istampar ta' Gutenberg. X'tista' tgħidilna din l-analoġija dwar il-futur tas-soċjetà tal-informazzjoni Ewropea u b'mod partikolari l-lingwi tagħna?

Wara l-invenzjoni ta' Gutenberg, skoperti reali fil-komunikazzjoni u skambju ta' għarfien kienu mwettqa permezz ta' sforzi bħat-traduzzjoni tal-Bibbja ta' Luther għal-lingwa komuni. Fis-sekli sussegwenti, tekniki kulturali ġew żviluppati biex jimmaniġġjaw aħjar l-ipproċessar tal-lingwa u l-iskambju ta' għarfien:

- l-istandardizzazzjoni ortografika u grammatikali ta' lingwi kbar ippermettiet it-tixrid rapidu ta' ideat xjentifiċi u intellettwali ġodda;
- l-iżvilupp ta' lingwi uffiċjali għamel possibbli għaċ-ċittadini li jikkomunikaw bejn ċerti konfini (ta' spiss politiċi);
- it-tagħlim u t-traduzzjoni tal-lingwi ppermettew skambju bejn il-lingwi;
- il-ħolqien ta' linji gwida ġurnalistiċi u biblijografiċi żguraw il-kwalità u d-disponibilità ta' materjal ipprintjat;
- il-ħolqien ta' midja differenti bħal gazzetti, radju, televiżjoni, kotba, u formati oħra ssodisfaw ħtiġijiet differenti ta' komunikazzjoni.

Fl-aħħar għoxrin sena, it-teknoloġija tal-informatika (TI) għenet biex ħafna mill-proċessi jiġu awtomatizzati u ffaċilitati:

- is-software ta' desktop publishing jissostitwixxi l-ittajpjar u t-typesetting;
- il-Microsoft PowerPoint tissostitwixxi t-trasparenzi tal-projectors;
- il-posta elettronika tibgħat u tirċievi dokumenti aktar malajr minn fax;
- Skype jagħmel telefonati bl-Internet u jospita laqgħat virtwali;
- il-formati ta' kodifikazzjoni audio u video jagħmluha faċli għal skambju ta' kontenut multimidjali;
- il-magni ta' tiftix jipprovdu aċċess ibbażat fuq kelma ewlenija għall-paġni web;
- is-servizzi fuq l-internet bħal Google Translate jipproduċu traduzzjonijiet ta' malajr u approssimattivi;
- il-pjattaformi tal-midja soċjali jiffaċilitaw il-kollaborazzjoni u l-qsim ta' informazzjoni.

Għalkemm għodod u applikazzjonijiet bħal dawn huma utli, bħalissa ma jistgħux jappoġġjaw b'mod suffiċjenti soċjetà tal-informazzjoni multilingwi Ewropea sostenibbli, soċjetà moderna u inklussiva fejn l-informazzjoni u l-merkanzija jistgħu jiċċirkolaw b'mod ħieles.

## Il-Konfini Lingwistiċi jfixklu s-Soċjetà Ewropea tal-Informazzjoni

Ma nistgħux inbassru b'mod preċiż kif se tkun is-soċjetà tal-informazzjoni tal-ġejjieni. Meta niġu biex niddiskutu strateġija għall-enerġija Ewropea jew politika barranija komuni, aħna nistgħu nkunu nixtiequ nisimgħu l-ministri tal-affarijiet barranin Ewropej jitkellmu bil-lingwa nattiva tagħhom. Aħna għandna mnejn nixtiequ pjattaforma fejn in-nies, li jitkellmu lingwi differenti u li għandhom profiċjenza diversa fil-lingwi, jistgħu jiddiskutu xi suġġett partikolari waqt li t-teknoloġija tiġbor l-opinjonijiet tagħhom awtomatikament u tiġġenera sommarji. Aħna nistgħu wkoll inkunu rridu nitkellmu ma' help desk ta' assigurazzjoni tas-saħħa li tinsab f'pajjiż barrani.

Huwa ċar li l-ħtiġijiet ta' komunikazzjoni għandhom kwalità differenti meta mqabbla ma' ftit snin ilu. F'ekonomija globali u spazju ta' informazzjoni, aktar lingwi, kelliema u kontenut jikkonfrontawna u jirrikjeduna li ninteraġixxu malajr ma' tipi ġodda ta' midja. Il-popolarità kurrenti ta' midja soċjali (Wikipedia, Facebook, Twitter u YouTube) hija biss parti żgħira minn stampa akbar.

*L-ekonomija globali u l-ispazju tal-informazzjoni jikkonfrontawna b'aktar lingwi, kelliema u kontenut.*

Illum, aħna nistgħu nittrażmettu gigabytes ta' test madwar id-dinja fi ftit sekondi qabel nagħrfu li huwa b'lingwa li aħna ma nifhmux. Skont rapport riċenti mitlub mill-Kummissjoni Ewropea, 57% tal-utenti tal-Internet fl-Ewropa jixtru oġġetti u servizzi b'lingwi li mhumiex il-lingwa nattiva tagħhom. (l-Ingliż huwa l-ilsien barrani l-aktar komuni segwit mill-Franċiż, il-Ġermaniż u l-Ispanjol.) 55% tal-utenti jaqraw kontenut f'lingwa barranija filwaqt li 35% biss jużaw lingwa oħra biex jiktbu ittri elettroniċi jew jibagħtu kummenti fuq il-web.[i] Ftit snin ilu, l-Ingliż seta' kien il-*lingua franca* tal-web—il-maġġoranza l-kbira tal-kontenut fuq il-web kien bl-Ingliż—iżda s-sitwazzjoni issa nbidlet drastikament. L-ammont ta' kontenut fuq l-internet b'lingwi oħra (partikolarment lingwi tal-Asja u Għarab) kiber f'daqqa waħda.

Firda diġitali li tinsab kullimkien u li hija kkawżata minn konfini lingwistiċi sorprendentement ma kisbitx ħafna attenzjoni fid-diskors pubbliku; madankollu, qajmet kwistjoni urġenti ħafna, "Liema lingwi Ewropej se jirnexxu u jippersistu fis-soċjetà tal-informazzjoni u l-għarfien, ibbażata fuq in-networks?"

*Liema lingwi Ewropej se jirnexxu u jippersistu fin-network ta' informazzjoni u s-soċjetà ta' għarfien?*

## Il-Lingwi Tagħna f'Riskju

L-istampar ikkontribwixxa għal skambju imprezzabbli ta' informazzjoni fl-Ewropa, iżda wassal ukoll għall-estinzjoni ta' bosta lingwi Ewropej. Lingwi reġjonali u dawk f'minoranza rarament ġew stampati. Bħala riżultat ta' dan, ħafna lingwi bħal Cornish jew Dalmatian spiss kienu limitati għal forom orali ta' trasmissjoni, li llimita l-adozzjoni kontinwa, it-tixrid u l-użu tagħhom.

Il-lingwi tal-Ewropa, madwar 60, huma wieħed mill-assi l-aktar prezzjużi tagħha u l-aktar importanti f'dak li huwa assi kulturali. In-numru kbir ta' lingwi Ewropej huwa wkoll parti vitali mis-suċċess soċjali tagħha. [ii] Filwaqt li l-lingwi popolari bħall-Ingliż jew l-Ispanjol żgur li se jżommu l-preżenza tagħhom fis-soċjetà diġitali u s-suq li qed jitfaċċaw, bosta lingwi Ewropej jistgħu jinqatgħu mill- komunikazzjonijiet diġitali u jsiru irrilevanti għas-soċjetà tal-Internet. Żviluppi bħal dawn żgur li ma jkunux mixtieqa. Minn naħa, opportunità strateġika tista' tintilef li jista' jkun iddgħajjef il-pożizzjoni globali tal-Ewropa. Min-naħa l-oħra, żviluppi bħal dawn jistgħu jmorru kontra l-għan ta' parteċipazzjoni ugwali għal kull ċittadin Ewropew irrispettivament mil-lingwa. Skont rapport tal-UNESCO dwar il-multilingwiżmu, il-lingwi huma mezz essenzjali għat-tgawdija tad-drittijiet fundamentali, bħall-espressjoni politika, l-edukazzjoni u l-parteċipazzjoni fis-soċjetà. [iii]

*Il-varjetà wiesgħa ta' lingwi fl-Ewropa hija waħda mill-assi kulturali l-aktar importanti u parti essenzjali mis-suċċess tal-Ewropa.*

## It-Teknoloġija Lingwistika hija Teknoloġija Katalizzanti Ewlenija

Fil-passat, l-isforzi tal-investiment iffukaw fuq l-edukazzjoni u t-traduzzjoni tal-lingwi. Pereżempju, skont ċerti stimi, is-suq Ewropew għat-traduzzjoni, l-interpretazzjoni, il-lokalizzazzjoni ta' software, u l-globalizzazzjoni ta' websajts kien ta' EUR 8.4 biljun fl-2008 u kien mistenni li jikber b'10% fis-sena. [iv] Madankollu, din il-kapaċità eżistenti mhix biżżejjed biex tissodisfa l-ħtiġijiet kurrenti u futuri.

It-teknoloġija lingwistika hija teknoloġija katalizzanti ewlenija li tista' tħares u trawwem lingwi Ewropej. It-teknoloġija lingwistika tgħin lin-nies jikkollaboraw, iwettqu negozju, jaqsmu l-għarfien u jieħdu sehem f'dibattiti soċjali u politiċi irrispettivament mill-ostakoli lingwistiċi jew il-ħiliet tal-kompjuter. It-teknoloġija lingwistika diġà qed tassisti kompiti ta' kuljum, bħal kitba ta' ittri elettroniċi, twettiq ta' tiftix fuq l-internet jew prenotazzjoni ta' titjiriet. Aħna nibbenefikaw minn teknoloġija lingwistika meta:

*It-teknoloġija lingwistika tgħin lin-nies jikkollaboraw, iwettqu negozju ,jaqsmu l-għarfien, u jieħdu sehem f'dibattiti soċjali u politiċi b'lingwi differenti.*

- nsibu informazzjoni permezz ta' magni ta' tiftix fuq l-Internet;
- niċċekkjaw l-ortografija u l-grammatika fil-word processor;
- naraw rakkomandazzjonijiet dwar prodotti f'ħanut online;
- nisimgħu struzzjonijiet verbali ta' sistema ta' navigazzjoni;
- nittraduċu paġni web permezz ta' servizz fuq l-internet.

It-teknoloġiji lingwistiċi deskritti f'dan id-dokument huma parti essenzjali minn applikazzjonijiet futuri innovattivi. It-teknoloġija lingwistika hija teknoloġija li tipikament tippermetti xogħol f'qafas ta' applikazzjoni akbar bħal sistema ta' navigazzjoni jew magna ta' tiftix. Dawn il-White Papers jiffokaw fuq il-prontezza ta' teknoloġiji ewlenin għal kull lingwa.

Fil-futur qrib, inkunu neħtieġu teknoloġija lingwistika għal-lingwi kollha Ewropej li tkun disponibbli, bi prezz raġonevoli u integrata sewwa f'ambjenti ta' software akbar. Esperjenza interattiva, multimidjali u multilingwi tal-utent mhijiex possibbli mingħajr teknoloġija lingwistika.

## Opportunitajiet għat-Teknoloġija Lingwistika

It-teknoloġija lingwistika tista' twettaq traduzzjoni awtomatika, produzzjoni ta' kontenut, ipproċessar ta' informazzjoni u ġestjoni ta' għarfien possibbli għal-lingwi Ewropej kollha. It-teknoloġija lingwistika tista' wkoll tkompli l-iżvilupp ta' interfaces intuwittivi bbażati fuq il-lingwa għal elettronika tad-dar, makkinarju, vetturi, kompjuters u robots. Għalkemm ħafna prototipi diġà jeżistu, l-applikazzjonijiet kummerċjali u industrijali għadhom fi stadji bikrin ta' żvilupp. Kisbiet riċenti fir-riċerka u l-iżvilupp ħolqu tieqa ġenwina ta' opportunità. Pereżempju, it-traduzzjoni awtomatika (TA) diġà qed tagħti ammont raġonevoli ta' preċiżjoni fi ħdan dominji speċifiċi, u applikazzjonijiet esperimentali jipprovdu informazzjoni multilingwi u ġestjoni ta' għarfien kif ukoll produzzjoni ta' kontenut f'ħafna ilsna Ewropej.

Applikazzjonijiet lingwistiċi, interfaces għall-utenti bbażati fuq il-vuċi u sistemi ta' djalogu jinsabu tradizzjonalment f'dominji speċjalizzati ħafna, u ħafna drabi jagħtu prestazzjoni limitata. Qasam wieħed attiv ta' riċerka huwa l-użu tat-teknoloġija lingwistika għal operazzjonijiet ta' salvataġġ f'zoni ta' diżastri. F' ambjenti bħal dawn ta' riskju għoli, l-eżattezza tat-traduzzjoni tista' tkun kwistjoni ta' ħajja jew mewt. L-istess raġunament japplika għall-użu tat-teknoloġija lingwistika fl-industrija tal-kura tas-saħħa. Robots intelliġenti b'kapaċitajiet lingwistiċi u bejn il-lingwi għandhom il-potenzjal li jsalvaw il-ħajjiet.

Hemm opportunitajiet tas-suq enormi fl-edukazzjoni u l-industriji tad-divertiment għall-integrazzjoni ta' teknoloġiji lingwistiċi fil-logħob, offerti ta' edudivertiment, ambjenti ta' simulazzjoni jew programmi ta' taħriġ. Servizzi mobbli ta' informazzjoni, software għat-tagħlim tal-lingwi assistit mill-kompjuter, ambjenti tal-eLearning, għodod għal awtoevalwazzjoni u software għal sejbien ta' plaġjariżmu huma biss ftit eżempji oħra fejn it-teknoloġija lingwistika jista' jkollha rwol importanti. Il-popolarità ta' applikazzjonijiet soċjali tal-midja bħal Twitter u Facebook jissuġġerixxu ħtieġa akbar għal teknoloġiji lingwistiċi sofistikati li jistgħu jissorveljaw postijiet, iwettqu sommarji ta' diskussjonijiet, jissuġġerixxu tendenzi ta' opinjonijiet, jiskopru reazzjonijiet emozzjonali, jidentifikaw ksur tad-drittijiet tal-awtur jew użu ħażin ta' sistemi ta' kompjuters.

*Il-multilingwiżmu huwa r-regola, mhux eċċezzjoni.*

It-teknoloġija lingwistika tirrappreżenta opportunità kbira għall-Unjoni Ewropea li tagħmel sens kemm ekonomikament kif ukoll kulturalment. Il-multilingwiżmu fl-Ewropa sar ir-regola. Negozji, organizzazzjonijiet u skejjel Ewropej huma wkoll multinazzjonali u diversi. Iċ-ċittadini jridu jikkomunikaw bejn il-konfini lingwistiċi li għadhom jeżistu fis-Suq Komuni Ewropew. It-teknoloġija lingwistika tista' tgħin biex jingħelbu dawn l-ostakli li fadal waqt li tappoġġja l-użu ħieles u miftuħ tal-lingwi. Barra minn hekk, teknoloġija lingwistika innovattiva u multilingwi għall-Ewropej tista' wkoll tgħinna nikkomunikaw mal-imsieħba globali tagħna u l-komunitajiet multilingwi tagħhom. It-teknoloġiji lingwistiċi jappoġġjaw l-ammont kbir ta' opportunitajiet ekonomiċi internazzjonali.

## Sfidi li t-Teknoloġija Lingwistika Taffaċċja

Għalkemm it-teknoloġija lingwistika għamlet progress konsiderevoli fl-aħħar ftit snin, il-pass preżenti tal-progress teknoloġiku u l-innovazzjoni tal-prodotti miexi bil-mod wisq. Ma nistgħux nistennew għaxra jew għoxrin sena biex jitwettaq titjib sinifikanti li jista' jżid il-komunikazzjoni u l-produttività fl-ambjent multilingwi tagħna.

*Il-pass preżenti tal-progress teknoloġiku huwa bil-mod wisq biex wieħed jasal għal prodotti ta' software sostanzjali fl-għaxra sal-għoxrin sena li ġejjin.*

Teknoloġiji lingwistiċi b'użu wiesa', bħal karatteristiċi ta' ortografija u grammatika fil-word processors, huma tipikament monolingwi, u huma disponibbli biss għal numru żgħir ta' lingwi. Applikazzjonijiet għal komunikazzjoni multilingwi jeħtieġu ċertu livell ta' sofistikazzjoni. Traduzzjoni awtomatika u servizzi fuq l-internet bħal Google Translate jew Bing Translator huma eċċellenti fil-ħolqien ta' approssimazzjoni tajba ta' kontenut f'dokument. Iżda servizzi fuq l-internet bħal dawn u applikazzjonijiet TA professjonali huma mimlija diffikultajiet varji meta jkunu meħtieġa traduzzjonijiet preċiżi ħafna u kompluti. Hemm ħafna eżempji magħrufa ta' traduzzjonijiet żbaljati u umoristiċi, pereżempju, traduzzjonijiet litterali tal-ismijiet *Bush* jew *Kohl*, li juru l-isfidi li t-teknoloġija lingwistika għad trid taffaċċja.

## Il-Ksib tal-Lingwi

Sabiex nuru kif il-kompjuters jimmaniġġjaw il-lingwa u għaliex il-ksib tal-lingwi huwa kompitu diffiċli ħafna, nagħtu ħarsa fil-qosor lejn il-mod kif il-bnedmin jiksbu l-ewwel u t-tieni lingwa, imbagħad nagħmlu skeċċ kif sistemi tat-traduzzjoni awtomatika jaħdmu—hemm raġuni għaliex il-qasam tat-teknoloġija lingwistika huwa marbut mill-qrib mal-qasam tal-intelliġenza artifiċjali.

Il-bnedmin jiksbu l-ħiliet lingwistiċi permezz ta' żewġ modi differenti. L-ewwel, it-tarbija titgħallem lingwa billi tisma' l-interazzjoni bejn il-kelliema tal-lingwa. Espożizzjoni għal eżempji konkreti lingwistiċi mill-utenti tal-lingwi, bħal ġenituri, aħwa u membri oħra tal-familja, jgħinu lit-trabi mill-età ta' madwar sentejn jew viċin jipproduċu l-ewwel kelmiet u frażijiet qosra tagħhom. Dan huwa possibbli biss minħabba d-dispożizzjoni ġenetika speċjali li għandhom il-bnedmin għat-tagħlim tal-lingwi.

*Il-bnedmin jiksbu ħiliet lingwistiċi permezz ta' żewġ modi differenti: tagħlim ta' eżempji u tagħlim tar-regoli bażiċi lingwistiċi.*

It-tagħlim tat-tieni lingwa normalment jeħtieġ sforz ferm aktar meta tifel jew tifla ma jkollhiex immersjoni fil-komunità lingwistika ta' kelliema nattivi. Fl-età skolastika, lingwi barranin jinkisbu ġeneralment permezz ta' tagħlim tal-istrutturi grammatikali, il-vokabularju u l-ortografija tagħhom minn kotba u materjal edukattiv li jiddeskrivu l-għarfien lingwistiku f'termini ta' regoli astratti, tabelli u testi bħala eżempji. It-tagħlim ta' lingwa barranija jieħu ħafna ħin u sforz, u jsir aktar diffiċli bl-età.

Iż-żewġ tipi ewlenin ta' sistemi ta' teknoloġija lingwistika jakkwistaw kapaċitajiet lingwistiċi b'mod simili bħall-bnedmin. Metodi statistiċi jiksbu għarfien lingwistiku minn ġbir vast ta' testi konkreti bħala eżempji f'lingwa waħda jew fl-hekk imsejħa testi paralleli li huma disponibbli f'żewġ lingwi jew aktar. L-algoritmi tat-tagħlim awtomatiku jiffurmaw ċerta fakultà lingwistika li tista' tikseb mudelli ta' kif kliem, frażijiet qosra u sentenzi sħaħ jintużaw b'mod korrett f'lingwa waħda jew jiġu tradotti minn lingwa għal

*Iż-żewġ tipi ewlenin ta' sistemi ta' teknoloġija lingwistika jiksbu l-lingwa b'mod simili bħall-bnedmin.*

oħra. In-numru kbir ta' sentenzi li metodi statistiċi jeħtieġu huwa enormi. Il-kwalità tal-prestazzjoni tiżdied hekk kif in-numru ta' testi analizzati jiżdiedu. Huwa komuni li sistemi bħal dawn jitħarrġu fuq testi li jinkludu miljuni ta' sentenzi. Din hija waħda mir-raġunijiet għaliex fornituri ta' magni ta' tiftix huma ħerqana li jiġbru kemm jista' jkun materjal bil-miktub. Il-korrezzjoni ortografika f'word processors, l-informazzjoni disponibbli fuq l-internet, u s-servizzi ta' traduzzjoni bħal Google Search u Google Translate jiddependu fuq metodu statistiku (mmexxi minn dejta).

Sistemi bbażati fuq regoli huma t-tieni tip ewlieni ta' teknoloġija lingwistika. Esperti mil-lingwistika, lingwistika kompjutazzjonali u x-xjenza tal-kompjuter jikkodifikaw analiżi grammatikali (regoli tat-traduzzjoni) u jikkumpilaw listi ta' vokabularju (dizzjunarji). It-twaqqif ta' sistema bbażata fuq regoli jieħu ħafna ħin u jinvolvi xogħol intensiv. Sistemi bbażati fuq regoli jeħtieġu wkoll esperti speċjalizzati sew. Uħud mis-sistemi ta' traduzzjoni awtomatika bbażati fuq regoli kienu taħt żvilupp kostanti għal aktar minn għoxrin sena. Il-vantaġġ tas-sistemi bbażati fuq regoli huwa li l-esperti jistgħu jikkontrollaw b'mod aktar dettaljat l-ipproċessar tal-lingwa. Dan jagħmel possibbli li l-iżbalji jiġu kkoreġuti b'mod sistematiku fis-software u tingħata informazzjoni dettaljata lill-utent, speċjalment meta s-sistemi bbażati fuq regoli jintużaw għat-tagħlim tal-lingwi. Minħabba restrizzjonijiet finanzjarji, teknoloġija lingwistika b'sistemi bbażati fuq regoli hija possibbli għal-lingwi ewlenin biss.

# Il-Malti fis-Soċjetà tal-Informazzjoni Ewropea[v]

## Fatti Ġenerali

Il-Malti huwa l-lingwa nazzjonali tal-arċipelagu Malti, li jikkonsisti fil-gżejjer ta' Malta, Għawdex u Kemmuna.

Flimkien mal-Ingliż, il-Malti huwa wkoll l-ilsien uffiċjali ta' Malta. Skont id-*Demographic Review 2009* mill-Uffiċċju Nazzjonali tal-Istatistika ta' Malta[vi], l-istima tal-popolazzjoni Maltija (minbarra l-barranin) fl-aħħar tas-sena 2009 kienet 396,278. Huwa stmat li llum, minħabba l-fażijiet tal-emigrazzjoni minn Malta l-aktar fil-ħamsinijiet u s-sittinijiet, bejn wieħed u ieħor l-istess numru ta' kelliema nattivi espatrijati jgħixu barra mill-pajjiż (l-aktar fir-Renju Unit, l-Awstralja, l-Istati Uniti u l-Kanada).

Għalkemm il-Malti jappartjeni għall-fergħa Għarbija tan-Nofsinhar tal-familja lingwistika Semitika, huwa pjuttost differenti mill-ilsna neo-Għarbin l-oħra. L-istruttura tiegħu huwa r-riżultat ta' sitwazzjonijiet ta' kuntatt lingwistiċi differenti li ffurmaw taħt mexxejja differenti tal-gżejjer fil-kors ta' millennju. Filwaqt li l-qalba tal-Malti hija Semitika, fiha wkoll superstrat Rumanz u adstrat Ingliż. Barra minn hekk, il-Malti huwa l-uniku lsien Semitiku miktub bl-alfabett Latin (modifikat).

Il-qalba Semitika tal-Malti tnisslet mill-konkwista Għarbija fit-870 AD u l-popolazzjoni mill-ġdid sussegwenti tagħha permezz ta' nies li ġew jgħixu f'Malta li jitkellmu bl-Għarbi. L-ewwel kuntatt dirett ma' lingwi Rumanzi ġie stabbilit fl-1090 meta Malta nħakmet minn Normanni, li ġabu l-Isqalli magħhom, filwaqt li l-popolazzjoni kienet għadha qed tuża l-Għarbi vernakulari tagħha fil-ħajja ta' kuljum. Malta kienet aktar u aktar maqtugħa politikament, kulturalment u lingwistikament mid-dinja Għarbija. Fis-sekli ta' wara, taħt l-influwenza tal-ilsna Rumanzi tal-mexxejja, aktar u aktar kliem Rumanz misluf daħal fid-djalett Għarbi. Meta Malta kienet taħt il-ħakma Ingliża fl-1800, il-lingwa uffiċjali nbidlet mit-Taljan għall-Ingliż, li ġab miegħu numru dejjem akbar ta' kliem Ingliż misluf fil-Malti[vii]. Il-sentenza li ġejja meħuda minn artiklu ta' gazzetta[viii] tista' turi l-influwenzi differenti tal-ilsna f'kuntatt (kliem Rumanz misluf huwa b'tipa grassa, kliem mill-Ingliż sottolinjat):

> *Il-<u>hold-up</u> sar minn żagħżugħ li kien liebes **nuċċali skur** tax-xemx.*
>
> *[The robbery was committed by a young man who was wearing dark sunglasses.]*

Wieħed mill-fatti notevoli dwar il-Malti huwa li minkejja n-numru relattivament żgħir tal-kelliema tiegħu u z-zona żgħira fejn hu mitkellem, hemm numru pjuttost rikk ta' varjanti jew djaletti. B'mod ġenerali, distinzjoni ewlenija tista' ssir bejn il-varjetà standard mitkellma fiz-zoni urbani bħall-Belt Valletta u tas-Sliema u varjetajiet mhux standard mitkellma fiz-zoni rurali. Barra minn Malta, il-Malti mitkellem fl-Awstralja żviluppa f'etnolett uniku msejjaħ *Maltraljan*. Dan huwa differenti mill-Malti Standard

prinċipalment f'dak li huwa l-lessiku tiegħu (jiġifieri, il-vokabularju) li huwa r-riżultat ta' kliem misluf b'mod estensiv mill-Ingliż (Awstraljan) u bidla sussegwenti fit-tifsira.

Minħabba li l-Ingliż huwa t-tieni ilsien uffiċjali f'Malta, ħafna Maltin huma bilingwi. Bejn il-pilastri tal-monolingwiżmu u l-bilingwiżmu sħiħ, hemm sekwenza ta' taħlit ta' lingwi u codeswitching. Fid-dar u bejniethom, ħafna Maltin jitkellmu biss bil-Malti. L-Ingliż, min-naħa l-oħra, huwa l-lingwa użata fil-kuntest ta' kitba f'edukazzjoni ogħla u fil-komunikazzjoni mal-barranin.

## Partikolaritajiet tal-Lingwa Maltija

Il-Malti huwa l-uniku lsien Semitiku fl-Unjoni Ewropea u l-uniku lsien Semitiku miktub bl-alfabett Latin. L-alfabett Malti jagħmel użu minn xi grafemi speċjali li jvarjaw minn alfabetti oħra Latini (il-valuri tal-ħoss huma mogħtija fl-Alfabett Fonetiku Internazzjonali): ċ [tʃ], ġ [dʒ], għ (il-biċċa l-kbira siekta), ħ [h], ż [z]

Xi karatteristiċi partikolari tal-Malti huma

- ordni tal-kliem ħielsa
- morfoloġija Semitika
- sistema temporali bbażata fuq l-aspett
- nuqqas ta' infinittiv morfoloġiku

Anki jekk ma fihx trufijiet tal-każi, il-Malti għandu ordni ta' kliem ħielsa ħafna. Is-sentenza *Il-kelb gidem il-qattusa lbieraħ* 'The dog bit the cat yesterday.' għandha l-ordni tal-kliem S(uġġett) V(erb) O(ġġett) iżda tista' wkoll tiġi espressa bħala:

| | | | | |
|---|---|---|---|---|
| Ilbieraħ | il-kelb | gidem | il-qattusa. | (SVO) |
| yesterday | the-dog (m) | he.bit | the-cat (f) | |

'Yesterday the dog bit the cat.'

| | | | | |
|---|---|---|---|---|
| Gidem | il-qattusa | l-kelb | ilbieraħ. | (VOS) |
| he.bit | the-cat(f) | the-dog | yesterday | |

'The dog bit the cat yesterday.'

| | | | | |
|---|---|---|---|---|
| Il-qattusa | gidimha | l-kelb | ilbieraħ. | (OVS) |
| the-cat(f) | he.bit.her | the-dog | yesterday | |

'The cat, it was bitten by the dog yesterday.'

Kif it-traduzzjonijiet bl-Ingliż jippruvaw juru, l-ordnijiet tal-kliem differenti għandhom enfasi differenti fit-tifsira. Fl-ewwel żewġ eżempji, l-ordni tal-kliem mhux immarka, bl-oġġett wara l-verb. Fl-aħħar eżempju, l-oġġett il-qattusa ('the cat') jippreċedi l-verb. Kif jissemma f'Fabri (1993: p. 140), din l-ordni tal-kliem hija mmarkata u tenfasizza l-oġġett għal kuntrast. Bl-oġġett quddiem, kelliema nattivi jippreferu jimmarkaw il-qattusa bl-enklitika tal-oġġett -ha mal-verb. Barra minn hekk, fid-diskors, dan il-kuntrast huwa espress b'intonazzjoni differenti. L-ordni tal-kliem fit-tieni eżempju (VOS) tista' tintuża biex tesprimi tifsira kuntrastiva kif ukoll, b'intonazzjoni xierqa, tqiegħed l-enfasi fuq *gidem il-qattusa* 'he bit the cat'. Mingħajr din it-tifsira kuntrastiva (u mingħajr l-

intonazzjoni kuntrastiva) l-enfasi tkun fuq il-fatt innifsu bħal: "Ma smajtx dak li ġara lbieraħ? Il-kelb gidem il-qattusa lbieraħ!" (Fabri, konverżazzjoni personali).

Bħala lsien Semitiku, il-Malti juri morfoloġija mhux konkatenattiva, jiġifieri inflettiva u l-forom ta' kliem imnisslin jinbidlu internament:

F'lingwi bħall-Ingliż, il-forom tal-kliem huma magħmula minn zkuk u affissi, jiġifieri b'mod konkatenattiv. Il-verb *shoot* jista' jkun ikkonjugat fit-terza persuna biż-żieda tal-affiss -s maz-zokk bħal f' *(he) shoot-s*. Barra minn hekk, miz-zokk verbali jista' jitnissel nom billi jiżdied l-affiss *-er* bħal f'*shoot-er*. Għaldaqstant kemm l-inflessjoni kif ukoll id-derivazzjoni jseħħu mingħajr tibdil fl-istruttura interna, jiġifieri b'mod konkatenattiv.

Fil-Malti, hemm taħlita ta' morfoloġija bbażata fuq iz-zokk morfemiku u morfoloġija bbażata fuq l-għerq u l-mudell. Fil-komponent Semitiku, l-"unità" bażika f'kelma ta' spiss ma tkunx iz-zokk iżda l-għerq magħmul minn tliet (xi kultant erba') konsonanti f'ordni fissa li ġġorr magħha tifsira ġenerali. Zkuk tal-kliem bit-tifsira speċifika tagħhom huma ffurmati billi l-konsonanti jiġu organizzati skont ċertu mudell. Pereżempju, l-għerq *k-t-b* iħaddan it-tifsira ta' dak kollu marbut mal-"kitba". F'dan li ġej, il-mudelli huma rappreżentati bħala numri **1,2,3** għall-konsonanti tal-għerq u **v** għall-vokali bejniethom, pereżempju **1v2v3**. Meta wieħed japplika l-mudell **1v2v3** u jimla l-pożizzjonijiet tal-vokali bejn il-konsonanti tal-għerq 1,2 u 3 bis-sekwenza tal-vokali i-e, wieħed jifforma l-verb *kiteb* 'he wrote'. L-inflessjoni ta' dan verb għall-plural issir bit-twaħħil tal-affiss tal-plural *-u*, li tagħti l-forma *kitbu* 'they wrote'. L-applikazzjoni tal-**1v22v:3** mal-għerq tagħti n-nom aġent *kittieb* 'writer'. L-inflessjoni tan-nom biż-żieda tal-affiss *-a* tagħti l-plural *kittieba* 'writers'. Wieħed jinnota li s-suffiss tal-plural *-a* jixbah lil markatur għall-femminili *-a* sabiex *kittieba* tista' wkoll tirreferi għal kittieb femminili. Is-suffissi l-oħra Semitiċi Maltin tal-plural huma *-in* bħal fi *mħallef* 'judge', *imħallfin* 'judges'; *-at/ -iet* bħal f'*kittieba* '(female) writer', *kittiebat* '(female) writers' ; *-ijiet* bħal fi *żmien* 'time', *żminijiet* 'times'.

Nomi fil-plural fil-Malti jistgħu jiġu ffurmati wkoll b'mod mhux konkatenattiv (l-hekk imsejħa forom ta' plural miksur), jiġifieri l-ebda affiss ma jiżdied, iżda n-nom jinbidel internament, pereżempju *ktieb* 'book' *vs. kotba* 'books'.

Verbi mislufa llum huma importati l-aktar permezz ta' klassi ta' verbi speċjali li tista' takkomoda zkuk mhux maħduma. Pereżempju, iz-zokk *park-* bl-Ingliż sar il-bażi tal-forom tal-verbi Maltin *pparkjajt, pparkjat, pparkja* 'I / she / he parked'. Illum, din il-klassi ta' verbi speċjali li qabel kienet klassi marġinali Semitika żdiedet fid-daqs minħabba l-influss ta' verbi mislufa mill-Ingliż. Dan huwa produttiv ferm, ħafna drabi iwassal għal self *ad-hoc* ta' verbi bl-Ingliż li diġà għandhom kontroparti Semitika bil-Malti. Pereżempju 'to download (a file)' tista' tiġi espress bl-użu tal-verb Semitiku *niżżel* (oriġinarjament tfisser 'he caused to come down'). Meta wieħed jieħu z-zokk Ingliż *download* u jimportah permezz tal-klassi tal-verb speċjali dan minflok jagħti forom bħal *ddawnlowdjajt, ddawnlowdjat, ddawnlowdja* 'I / she / he downloaded'. Din l-istrateġija tiġi ta' spiss ikkritikata li qed tikkorrompi l-lingwa (Fabri, dalwaqt jiġi ppubblikat[a]: p. 17).

Il-verbi bil-Malti huma aċċentwati għall-aspett, jiġifieri jekk azzjoni tkunx kompluta (perfettiva) jew mhux kompluta (mhux perfettiva [ix]). Fin-nuqqas ta' kwalunkwe markaturi grammatikali oħra, verbi perfettivi huma interpretati bħala 'it-temp tal-passat' u verbi li mhumiex perfettivi bħala 'it-temp tal-preżent': *Andrew kiteb* 'Andrew wrote'; *Andrew jikteb* 'Andrew writes'. Il-kumbinazzjoni tal-verb li mhuwiex perfettiv ma' *kien*, tesprimi passat abitwali: *Andrew kien jikteb* 'Andrew used to write'. Iż-żieda tal-kelma *qed* 'progressiva' (bħall-forma bl-Ingliż *-ing*) tagħti *Andrew kien qed jikteb* 'Andrew was writing' eċċ.

Il-verbi Maltin ma għandhomx infinittivi morfoloġiċi. Għaldaqstant, fi predikati kumplessi bħal fis-sentenza bl-Ingliż 'Andrew wants to write', iż-żewġ verbi huma morfoloġikament finiti: *Andrew jrid jikteb* (litteralment: 'Andrew he wants he writes') anki jekk min-naħa semantika, *jikteb* mhuwiex finit.

## Żviluppi riċenti

Bl-avvanz tal-Ingliż għal-lingwa internazzjonali u l-lingwa tat-teknoloġija wara t-Tieni Gwerra Dinjija, l-ammont ta' kliem misluf mill-Ingliż fil-Malti kiber b'mod sostanzjali. Ħafna minnhom saru "nattivi", jiġifieri ġew adottati fl-użu regolari tant li anki kliem meħud mis-Semitiku ma jistax jieħu posthom. Pereżempju, minflok il-kelma użata ta' spiss *ajruport* (mill-Ingliż *airport*), il-kelma Semitika *mitjar* kienet proposta (imnissla minn *tar* 'he flew'). Madankollu, din qatt ma ġiet aċċettata mill-komunità lingwistika. Min-naħa l-oħra, kliem misluf jidħol fil-lingwa pjuttost malajr, jiġi impurtat b'mod spontanju, anki jekk diġà hemm kliem Malti tajjeb għalihom (pereżempju *ddawnlowdja vs niżżel* 'he downloaded'). Dan iqabbad biżgħat fost xi wħud li l-lingwa tista' ssir "korrotta" (Fabri, dalwaqt jiġi ppubblikat[a]: p. 17).

Żvilupp ieħor riċenti għall-Malti huwa l-istatus tiegħu bħala lingwa uffiċjali tal-Unjoni Ewropea. Dan għandu kemm vantaġġi, kif ukoll żvantaġġi (Fabri, dalwaqt jiġi ppubblikat[a]: p. 20). Minn naħa waħda, il-Malti finalment sar ilsien rikonoxxut internazzjonalment, status li ma kellux għal żmien twil, kien imwarrab bħala l-"lingwa tal-kċina" fis-sekli ta' qabel. Min-naħa l-oħra, it-tradutturi tal-UE Maltin qed jaffaċċjaw ċerti sfidi: bosta termini tekniċi u legali għad iridu jiġu "ivvintati" bil-Malti. Dan jirriżulta eventwalment f'espansjoni lessikali tal-lingwa (tabilħaqq aspett pożittiv), li, madankollu, għandu jiġi kkoordinat minn korpus ċentrali sabiex tradutturi individwali ma joħolqux termini differenti għall-istess kunċetti indipendentement minn xulxin (li hija problema serja). Il-korp ċentrali biex jittratta din l-isfida huwa l-Kunsill Nazzjonali għall-Ilsien Malti (*National Council for the Maltese Language*).

Żviluppi oħra fis-snin riċenti jikkonċernaw l-ortografija Maltija. Il-Malti (flimkien mal-Ingliż) sar l-ilsien uffiċjali ta' Malta fl-1 ta' Jannar, 1934 – bl-ortografija maħruġa mill-Għaqda tal-Kittieba tal-Malti (*Union of Maltese Writers*) fl-1924. Minn dakinhar, l-ortografija għaddiet minn tliet reviżjonijiet (1984, 1992 u 2008).

L-aħħar riforma ġiet rilaxxata fl-2008. L-għan tagħha kien li jitnaqqsu l-inċertezzi tal-kittieba li jirriżultaw minn numru konsiderevoli ta' varjanti ortografiċi għal ċerti kliem. Kif id-dokument *Deċiżjonijiet 1* tal-Kunsill (2008[a]) jirrimarka, ammont kbir ta' varjanti jista' jitnaqqas billi jinstab bilanċ konsistenti bejn l-ortografija grammatikali u fonetika. Għaldaqstant l-erba' varjanti

*zobtu, zoptu, sobtu* u *soptu* ('suddenly, unexpectedly') jistgħu jitnaqqsu għal żewġ varjanti għal *zoptu* ['zɔp.tʊ] u *soptu* ['sɔp.tʊ]. Għal raġuni simili, il-kelma *skond* [skɔnt] 'according to' nbidlet għal *skont* minħabba li l-forom grammatikali tagħha l-oħra ma jiġġustifikawx ortografija b' *d* (imnissla minn *secondo* Taljan), bħal pereżempju *skontok* ['skɔn.tɔk] 'according to you'.

Għat-tielet qasam (tal-kliem misluf), il-prinċipju jibqa' li l-kliem misluf jinkiteb skont l-ortografija Maltija jekk dawn jitqiesu bħala "nattivi" u jekk dan ma joħloqx kunflitti fil-pronunzja jew ma' regoli oħra tal-kitba Maltija. Madankollu, ħafna Maltin jippreferu li jiktbu kliem Ingliż misluf bl-ortografija oriġinali tagħhom, minħabba li drawhom. Fil-fatt, waqt seminar pubbliku dwar l-użu ta' kliem Ingliż misluf f'April 2008, kien hemm diskussjonijiet emozzjonali fost l-udjenza fil-każ ta' kliem bħal *email* u l-ortografija l-ġdida tagħha propost bħala *imejl.* Fatturi bħal drawwiet tal-komunità lingwistika jagħmlu l-istandardizzazzjoni tal-ortografija saħansitra aktar diffiċli milli li jinstab bilanċ bejn il-prinċipji grammatikali u fonetiċi.

Dawn l-eżempji jagħtu biss idea żgħira tal-ħidma iebsa li l-Kunsill Nazzjonali għall-Ilsien Malti qed iwettaq bħala parti mill-kultivazzjoni tal-lingwa f'Malta. Is-sezzjoni li jmiss se tagħti ħarsa lejn l-istorja tal-kultivazzjoni tal-lingwa f'Malta.

## Il-Kultivazzjoni tal-Lingwa f'Malta

Meta mqabbel ma' lingwi oħra tal-Ewropa, l-istatus tal-Malti bħala lsien uffiċjali (mill-1934) fih innifsu huwa żvilupp riċenti. Għaldaqstant il-kultivazzjoni tal-lingwa wkoll bdiet tard.

Għal sekli sħaħ, il-Malti kien biss il-mezz mitkellem tal-popolazzjoni Maltija u kien imwarrab meta mqabbel mal-lingwa uffiċjali rispettiva tal-mexxejja ta' Malta. Dan beda jinbidel mal-moviment tal-lingwa ta' nofs-/tmiem is-seklu 18 meta l-ewwel studji sistematiċi lingwistiċi tmexxew minn Agius de Soldanis (1750) u Mikiel Anton Vassalli (1797). Speċjalment Vassalli ppromwova l-ilsien Malti billi nkoraġġixxa l-użu tiegħu f'kull qasam tal-ħajja ta' kuljum. It-traduzzjonijiet tal-bibbja ta' Fortunato Panzavecchia f'nofs is-seklu 19 ikkontribwixxew għal aktar standardizzazzjoni tal-lingwa. Barra minn hekk, permezz tal-pass lejn ortografija standardizzata fil-bidu tas-seklu 20, ittieħed pass importanti mill-fondazzjoni tal-Għaqda tal-Kittieba tal-Malti (*Union of Maltese Writers*) fl-1920. Is-sistema ortografika, li ġiet żviluppata minn din l-organizzazzjoni, saret l-ortografija uffiċjali ta' Malta fl-1934 u, b'xi bidliet u żidiet, ilha tintuża minn dak iż-żmien.

Fl-1964, wara li nkisbet l-indipendenza mill-Gran Brittanja, l-istatus tal-Malti bħala lingwa nazzjonali u bħala lingwa uffiċjali flimkien mal-Ingliż inkiteb fil-kostituzzjoni. Meta Malta ssieħbet fl-UE fl-2004, il-Malti sar ilsien uffiċjali tal-UE. Kif issemma fit-taqsima t'hawn fuq, dan wassal għal ċerti sfidi, li jistgħu jissolvew biss minn korpus li jikkoordina l-istandardizzazzjoni u l-prassi komuni fix-xogħol tat-traduzzjoni.

Il-korp f'Malta biex jagħmel dan ix-xogħol huwa l-Kunsill Nazzjonali għall-Ilsien Malti (*National Council for the Maltese*

*Language*). Dan twaqqaf fl-2005 bħala l-ewwel organizzazzjoni tal-gvern sabiex tittratta uffiċjalment kwistjonijiet lingwistiċi u ppjanar lingwistiku għal-lingwa Maltija. Il-kompiti tal-Kunsill huma, kif imniżżlin fl-Att tal-Ilsien Malti (ATT Nru V tal-2004): il-jippromwovi l-ilsien Malti, "jadotta politika, pjan u strateġija lingwistika xierqa" u jwettaq dan fil-prattika. Xogħol ieħor importanti tal-Kunsill huwa li jaġġorna l-ortografija Maltija u jiddeċiedi fuq ortografija korretta (jieħu f'idejh l-inkarigu mill-Akkademja tal-Malti u b'hekk ikun prinċipalment responsabbli għar-riforma għall-ortografija Maltija tal-2008). Fuq is-sitweb tiegħu, il-Kunsill joffri wkoll korsijiet ta' taħriġ għall-qarrejja tal-provi u korsijiet tal-lingwa Maltija għall-barranin[x].

Qabel twaqqaf il-Kunsill, l-istandardizzazzjoni tal-ortografija kienet il-kompitu tal-Akkademja tal-Malti (*Academy of Maltese*). Din oriġinat fl-1964 mill-Għaqda tal-Kittieba tal-Malti (*Union of Maltese Writers*), li kienet il-korp li waqqaf l-ewwel ortografija uffiċjali fl-1924 / 1932. Illum l-għan prinċipali tal-Akkademja huwa li tippromwovi studji akkademiċi fil-lingwa u l-letteratura Maltija, tippromwovi l-użu tal-Malti f'kull qasam tal-ħajja ta' kuljum u tibni kuntatti mal-persuni li huma ħbieb tal-lingwa u li jużawha barra minn Malta[xi]. L-Akkademja taħdem mill-qrib flimkien mal-Kunsill Nazzjonali għall-Ilsien Malti.

Il-motivazzjoni wara l-Att tal-Ilsien Malti kienet l-idea li lingwa nazzjonali waħda li hija kondiviża mill-individwi kollha fi ħdan dak in-nazzjon tifforma l-bażi għall-identità kulturali u nazzjonali. Dan naturalment jeħtieġ standardizzazzjoni tal-lingwa. Fil-fatt, mill-moviment tal-kultivazzjoni tal-lingwa mis-seklu 19 sal-lum, il-Malti avvanza minn ilsien imwarrab u vernakulari kif kien qabel għal ilsien nazzjonali ta' prestiġju għoli. Dan jidher ukoll fl-ammont dejjem jikber ta' xogħlijiet letterarji bil-Malti matul l-istess perjodu ta' żmien u fin-numru kbir ta' organizzazzjonijiet influwenti u l-korpi għal-lingwa u l-letteratura Maltija (ara Fabri, dalwaqt jiġi ppubblikat[a]: p. 22).

## Il-Lingwi fl-Edukazzjoni

Partikolarment f'soċjetà bilingwali bħal dik ta' Malta, diversi aspetti għandhom rwol meta niġu għal-lingwa fl-edukazzjoni.

Aspett wieħed huwa l-lingwa ta' istruzzjoni, jiġifieri l-lingwa li tintuża uffiċjalment mill-għalliema matul il-lezzjonijiet fl-iskola jew fis-seminars fl-università.

Fattur ieħor huwa l-lingwa użata f'ċerti kotba tal-iskola. Bl-Ingliż bħala l-lingwa tax-xjenzi teknoloġiċi u naturali, ħafna mill-kotba tal-iskola dwar dawn is-suġġetti huma bl-Ingliż. Fil-fatt, l-isforzi biex jiġu tradotti termini tekniċi u xjentifiċi għall-Malti ltaqgħu ma' bosta problemi, waħda minnhom hija l-aċċettazzjoni mill-komunità tal-lingwa. Għaldaqstant is-suġġetti skolastiċi, ukoll, possibbilment jiddeterminaw il-lingwa ta' istruzzjoni għal ċerti lezzjonijiet, għalkemm jista' jkun ukoll li l-kotba tal-iskola bl-Ingliż (u t-terminoloġija bl-Ingliż li tinsab fihom) jintużaw waqt li l-lingwa tat-tagħlim tkun il-Malti.

Madankollu aspett ieħor huwa l-lingwa użata mill-individwi. Kelliema bilingwi mhux biss jużaw lingwi differenti f'kuntesti soċjali differenti ("dominji"), eż. il-Malti ma' tal-familja fid-dar, l-Ingliż mal-barranin, il-Malti jew l-Ingliż matul il-lezzjonijiet tal-

iskola eċċ. Dawn għandhom tendenza wkoll li jużaw iż-żewġ lingwi flimkien, jew iħalltu ż-żewġ lingwi (eż. kliem bl-Ingliż jitħalltu f'konverżazzjoni li tkun qed issir bil-Malti) jew permezz tal-codeswitching (eż. konverżazzjoni bil-Malti tinqaleb għall-Ingliż u lura għall-Malti, bil-partijiet tal-Ingliż ikunu akbar minn kliem biss waħedhom, iżda spiss jikkonsistu minn diversi sentenzi). Għaldaqstant anki matul il-lezzjonijiet tal-iskola li jiġu mgħallma b'lingwa waħda, il-konverżazzjonijiet bejn l-għalliema u l-istudenti jistgħu jaqilbu bejn il-lingwi (Camillieri, 1995).

Meta wieħed iżomm dawn it-tliet fatturi f'moħħu, wieħed jinduna li l-espożizzjoni attwali tal-istudenti għal-lingwa rispettiva fl-iskejjel jew fl-università hija xi ħaġa differenti mil-lingwa magħżula ta' istruzzjoni.

Rigward il-lingwa uffiċjali ta' istruzzjoni fl-edukazzjoni, kemm il-Malti kif ukoll l-Ingliż jintużaw fl-iskejjel u fl-università, minħabba li l-Malti u l-Ingliż jaqsmu l-istatus bħala lingwi uffiċjali ta' Malta. Fl-iskejjel, it-tnejn li huma jiġu mgħallma bħala suġġetti minn età bikrija. Liema lingwa tintuża bħala l-lingwa ta' istruzzjoni jiddependi mit-tip ta' skola. Skejjel privati għandhom it-tendenza li jużaw l-Ingliż aktar mill-Malti (xi kultant b'mod aktar estensiv), filwaqt li fl-iskejjel Maltin tal-istat il-Malti huwa kemxejn preferut mill-Ingliż. L-iskejjel tal-knisja għandhom il-preferenzi individwali tagħhom jiġifieri li xi wħud tradizzjonalment jippreferu lingwa waħda minn oħra.

Kif issemma qabel, il-biċċa l-kbira tal-kotba tax-xjenza li jintużaw fl-iskola huma bl-Ingliż. Għaldaqstant, bl-introduzzjoni ta' aktar u aktar suġġetti xjentifiċi aktar 'il quddiem fl-iskola u iżjed fl-università, l-istudenti huma esposti għal żewġ lingwi fl-istess ħin, li jintużaw f'sitwazzjonijiet differenti: jista' jkollhom il-lezzjonijiet tagħhom mgħallma bil-Malti, iżda jaqraw il-kotba tagħhom u jiktbu l-essays tagħhom bl-Ingliż. Speċjalment għal studenti tal-università, il-konverżazzjonijiet bejniethom, mal-ħbieb u l-lecturers spiss iseħħu bil-Malti, xi kultant jużaw il-code-switching /iħalltu bejn il-Malti jew ikunu saħansitra bl-Ingliż biss (tal-aħħar pereżempju ma' studenti internazzjonali jew lecturers).

Madankollu, fid-dar mal-familja tagħhom u l-ħbieb, ħafna Maltin jitkellmu bil-Malti, xi wħud iħalltu l-lingwi u ftit familji biss jitkellmu bl-Ingliż biss.

Kif jidher mill-eżempji ta' hawn fuq, minkejja l-fatt li kemm il-Malti kif ukoll l-Ingliż jintużaw bħala lingwi fl-edukazzjoni, hemm distribuzzjoni ċara meta niġu għall-użu tagħhom fis-soċjetà. Sciriha u Vassallo (2001, p. 29, iċċitati f'Fabri, dalwaqt jiġi ppubblikat[a]: p. 18) isemmu li "70% ta' dawk li wieġbu qalu li jużaw il-Malti fuq ix-xogħol, filwaqt li 90% qalu li jikkomunikaw mal-membri tal-familja tagħhom fid-dar bil-Malti. ... il-persentaġġi għall-Malti mitkellem huma għolja ħafna iżda jonqsu f'ħiliet oħra bħall-qari u l-kitba."

Din id-distribuzzjoni tal-Malti li qed jintuża prinċipalment bħala l-mezz mitkellem u l-Ingliż bħala l-mezz tal-kitba toħloq ċertu riskju, minħabba li jista' jkollha impatt fuq il-ħiliet differenti tal-kelliema nattiva tagħha f'dak li għandu x'jaqsam ma' taħdit, qari jew kitba. Sabiex wieħed jagħti r-raġunijiet għal dan il-fatt, wieħed għandu jħares lejn il-karatteristiċi bażiċi tal-lingwa mitħaddta u dik miktuba.

B'mod ġenerali, it-testi miktubin jiddistingwu ruħhom mid-diskors f'numru ta' modi. Li għandhom komuni huwa li t-tnejn huma modi ta' trasferiment ta' informazzjoni bejn iż-żewġ partijiet, jiġifieri l-kelliem u min qed jisma', u l-kittieb u l-qarrej, rispettivament. Madankollu, huma differenti fil-mod kif l-informazzjoni tgħaddi bejniethom. Fi kliem sempliċi, test miktub, kuntrarju għal diskors, iseħħ barra minn sitwazzjoni komunikattiva, interattiva u konkreta. Minn naħa, id-diskors jiddependi fuq l-interazzjoni bejn il-kelliem u min qed jisma'. Il-kelliem irid jibni struttura tal-informazzjoni b'ċertu mod. Dan huwa importanti minħabba l-memorja limitata u qasira tal-bniedem: min qed jisma' fil-konverżazzjoni jista' jassorbi ċertu ammont ta' informazzjoni biss qabel ma jkollu jinterrompi u jsaqsi lill-kelliem biex jiżgura li fehem.

Test miktub, min-naħa l-oħra, mhuwiex interattiv safejn il-qarrej ma' jistax jitlob għal aktar informazzjoni speċifika. Madankollu huwa jista' jara x'hemm 'il quddiem u lura fit-test (xi ħaġa li min qed jisma' ma jistax jagħmilha fid-diskors). B'dan il-mod, it-test innifisu miktub iservi bħala memorja fit-tul għall-qarrej. Għaldaqstant, test miktub jistruttura l-informazzjoni b'mod differenti minn kif isir f'konverżazzjoni mitħaddta. Pereżempju, test għandu jipprovdi aktar informazzjoni ta' sfond sabiex jagħti bażi komuni lill-qarrej qabel tibda għaddejja l-informazzjoni attwali. Din ma tkunx problema, jekk it-test jista' jservi bħala memorja fit-tul għall-qarrej. Fil-fatt, dan jippermetti struttura aktar elaborata mid-diskors, jiġifieri normalment ikun fih sentenzi itwal u ammont ogħla ta' propożizzjonijiet subordinati.

Din id-distinzjoni tar-reġistru (jiġifieri "l-istil tal-lingwa") hija dik li fil-letteratura (eż. Biber, 1991) issejħet strutturi ta' testi *orali versus letterati*. Tabilħaqq, test jista' jkun miktub f'reġistru orali li jixbah konverżazzjonijiet mitħaddta (eż. f'forum ta' iċċettjar jew posta elettronika informali). Iżda dan mhuwiex ir-reġistru normalment użat pereżempju f'essays. Idealment, il-kelliema nattivi jiksbu r-reġistru letterat diġà minn età żgħira, eż. permezz tal-ġenituri tagħhom li jaqrawlhom l-istejjer. Aktar tard fl-iskola, dan l-għarfien jissaħħaħ minn, pereżempju, eżerċizzju attiv tal-kitba tal-essays.

Reġistru letterat jiżviluppa maż-żmien f'lingwa bi tradizzjoni letterarja. Il-Malti, meta mqabbel mal-istorja qasira tiegħu bħala lsien uffiċjali miktub (mill-1934) għandu storja letterarja twila u rikka. Anki jekk l-eqdem letteratura skoperta hija skarsa ħafna (*Il Cantilena* minn Pietro Caxaro, li tmur lura għal madwar l-1450), tradizzjoni letterarja bdiet tifforma madwar l-erbgħinijiet fis-seklu 17. Fis-seklu 19, l-ammont ta' letteratura bil-Malti kienet qed tikber (Fabri, dalwaqt jiġi ppubblikat[a]: p. 25), u flimkien magħha, il-Malti kien qed jespandi. Illum huwa lsien li għandu reġistru letterat komplut.

Madankollu, dan ir-reġistru, jeħtieġ li jiġi pprattikat sabiex jinżamm l-istatus tal-lingwa bħala lingwa kemm konverżazzjonali kif ukoll letterarja. It-tendenza fl-edukazzjoni ogħla biex jinkitbu essays aktar bl-Ingliż milli bil-Malti, mill-anqas teoretikament, toħloq ir-riskju li l-Malti jibqa' jintuża f'reġistru orali biss. Ammont ogħla ta' websajts Maltin tal-ġeneri kollha huwa mixtieq biex wieħed ikopri ż-żewġ reġistri u s-sottotipi tagħhom u jiġi żgurat status stabbli tal-lingwa fir-rikkezza kollha tagħha.

## Aspetti internazzjonali

Meta wieħed iżomm f'moħħu t-taqsimiet preċedenti, issa għandu jkun mifhum li l-aspetti internazzjonali tal-Malti huma pjuttost differenti minn lingwi oħra. B'anqas minn miljun kelliema nattivi madwar id-dinja, il-Malti huwa kkunsidrat bħala lingwa "mitkellma anqas". Fl-istorja tiegħu, il-Malti ma kienx l-ilsien tal-okkupanti iżda wieħed ta' dawk li qed jokkupaw il-post. Bħala riżultat ta' dan, il-Malti qatt ma kien meqjus bħala lingwa internazzjonali jew lingua franca kif kien il-każ eż. tal-Latin, l-Ispanjol, il-Portugiż jew l-Ingliż, li kollha huma l-lingwi tal-konkwistaturi. Il-Malti tabilħaqq infirex lejn pajjiżi oħra, fejn għadu mitkellem sal-lum (l-Awstralja, il-Kanada, l-Istati Uniti u r-Renju Unit), iżda bħala lingwa tal-komunità biss. Kien jeħtieġ kważi 200 sena mill-ewwel interess tal-grammatiċi Maltin fil-lingwa tagħhom sakemm eventwalment kiseb l-istatus ta' lingwa uffiċjali. Saħansitra dakinhar, il-lingwa uffiċjali l-oħra, l-Ingliż, serviet bħala l-lingwa għar-relazzjonijiet internazzjonali.

Il-bidla biex il-Malti jsir ilsien internazzjonalment viżibbli seħħet mas-sħubija ta' Malta fl-UE fl-2004. Minn dakinhar, il-Malti huwa lsien uffiċjali fl-Unjoni Ewropea, flimkien bil-benefiċċji u l-isfidi kollha li huma marbutin ma' dan l-istatus.

Akkademikament, l-interess fil-Malti bħala suġġett tax-xjenza jmur lura sal-1603 meta Hieronymus Megiser ippubblika t-*Thesaurus Polyglottus* tiegħu, li kien jinkludi lista ta' kliem bil-Malti. L-ewwel studjuż li b'mod sistematiku esplora u ppromwova l-lingwa Maltija kien Mikiel Anton Vassalli. Huwa ppubblika grammatika (1790), dizzjunarju (1797) u alfabetti diversi (1788 u 1790) għall-Malti u llum huwa msejjaħ "il-missier tal-Ilsien Malti" (Brincat 2011). Fis-seklu 20, kien ippubblikat il-*Grammar of the Maltese Language* (1936) ta' Sutcliffe. Mis-sittinijiet tas-seklu 20, il-Lingwistika tal-Ilsien Malti kisbet għarfien akkademiku internazzjonali permezz tal-pubblikazzjonijiet ta' Joseph Aquilina (eż. *Papers in Maltese Linguistics* (1961) u *Maltese-English Dictionary,* żewġ volumi (1987 and 1990)). Minn dak iż-żmien, aktar u aktar studjużi barra minn Malta wrew interess fil-Malti. L-2007 rat it-twaqqif tal-Għaqda Internazzjonali tal-Lingwistika Maltija (*International Association of Maltese Linguistics*)[xii], assoċjazzjoni ta' lingwisti li huma interessati fil-lingwa Maltija. L-għan ewlieni tal-GĦILM, kif jidher fuq is-sitweb tagħha, huwa li tipprovdi "konnessjoni bejn studjużi interessi fil li ġejjin mid-dixxiplina kollha tal-Lingwistika", b'hekk tiffaċilita r-riċerka dwar il-Malti. Din l-għaqda trid ukoll li tgħaqqad flimkien nies minn sfondi differenti li jaħdmu bl-ilsien Malti (lingwisti, tradutturi, studenti u oħrajn).

## Il-Malti fuq l-Internet

Stħarriġ tal-Uffiċċju Nazzjonali tal-Istatistika ta' Malta fit-tieni kwart tal-2009[xiii] juri li fost il-popolazzjoni ta' madwar 400,000 ruħ, 67 fil-mija kellhom aċċess għall-kompjuter u 64 fil-mija kellhom aċċess għall-internet. Stħarriġ riċenti tal-Ewrobarometru (ippubblikat f'Mejju tal-2011)[xiv] dwar id-drawwiet tat-tiftix fuq l-internet fost l-utenti Ewropej wera li huma biss 6.5 fil-mija tal-utenti tal-internet Maltin li jużaw esklussivament il-Malti fuq l-internet meta jaqraw, jikkunsmaw il-kontenut jew jikkomunikaw. Minflok, 90.6 fil-mija jagħżlu li jibbrawżjaw is-websajts bl-Ingliż u 20.1 fil-mija bit-Taljan, rispettivament. Dawn iċ-ċifri iffurmaw il-bażi tal-artiklu fil-gazzetta Maltija li toħroġ kuljum "The Times of

Malta", li ħoloq diskussjoni interessanti l-aktar fost il-qarrejja Maltin tal-edizzjoni fuq l-internet[xv].

Madankollu, ir-riżultati eżatti tal-istħarriġ, iwasslu għall-konklużjoni li din l-abitudni mhix għażla maħsuba:

Meta mistoqsija liema lingwa l-Maltin jikkunsidraw bħala l-lingwa materna tagħhom, 89.5 fil-mija ta' dawk li wieġbu qalu li l-Malti kien l-ilsien nattiv tagħhom (meta mqabbel ma' 7.6 fil-mija biss għall-Ingliż u 0.2 fil-mija għat-Taljan).

Il-lingwi l-oħra barra dik użata minn dawk li wieġbu biex jaqraw jew jaraw kontenut fuq l-internet kienu l-Ingliż (90.6 fil-mija) u t-Taljan (20.1 fil-mija). 6.5 fil-mija biss wieġbu li jużaw il-lingwa tagħhom, li mhuwiex fatt sorprendenti, minħabba li ħafna Maltin huma bilingwali bil-Malti u bl-Ingliż u numru konsiderevoli jitkellmu bit-Taljan ukoll.

F'dak li għandu x'jaqsam ma' kitba fuq l-internet, in-numri favur il-Malti huma ogħla meta l-utenti jaqraw jew jaraw kontenut: 87 fil-mija qalu li jużaw il-Malti, 85 fil-mija l-Ingliż u 8 fil-mija t-Taljan.

Ir-raġuni li l-maġġoranza jużaw l-Ingliż bħala l-lingwa biex jikkunsmaw kontenut fuq l-internet tista' tkun sempliċiment in-numru limitat ta' websajts bil-Malti minflok il-preferenza għall-Ingliż fiha nnifisha. Niftakru li ħafna minn dawk li wieġbu ma jikkunsidrawx l-Ingliż bħala l-lingwa tagħhom u li l-użu tal-Malti żdied meta ġie prodott kontenut fuq il-web, anki jekk dan l-użu tal-Malti fil-maġġoranza tal-każijiet iseħħ f'forums ta' iċċettjar u pjattaformi soċjali, għalhekk fi stil ta' lingwaġġ tat-taħdit, jiġifieri fir-reġistru orali.

Karatteristika partikolari dwar il-Malti użat mill-ġenerazzjoni żagħżugħa fi pjattaformi soċjali u forums ta' iċċettjar hija l-ortografija fonetika tagħha, mingħajr il-karattri bħal *għ* siekta u l-*h*. Għaldaqstant *għax* 'because' tinkiteb *ax*, *tiegħi* 'my' *tiei* eċċ. Ir-raġuni għal dan tista' tkun l-introduzzjoni tard tal-karattri speċjali tal-Malti fid-dinja tal-PC. Minkejja li l-Malti ġie implimentat fil-qafas tal-Unicode mill-bidu tiegħu, kompjuters u sistemi ta' operazzjoni segwew ħafna aktar tard. L-Awtorita Maltija tal-Istandards ħarġet forma standardizzata ta' tastiera Maltija fl-2002, u s-sistema ta' operazzjoni Windows tal-Microsoft kienet disponibbli fil-verżjoni tal-lingwa Maltija fl-2006 biss (mal-Windows XP). Fil-każ ta' telefonijiet ċellulari, l-ittri speċjali Maltin għadhom ma ġewx implimentati. Għaldaqstant naraw jekk l-ortografija *ad hoc* tal-forums tal-iċċettjar hix se twitti t-triq għal ortografija b'karattri speċjali ladarba dawn ikunu disponibbli fuq it-telefonijiet ċellulari jew jekk din l-ortografija fonetika se tkompli teżisti bħala "soċjolekt" tal-ġenerazzjoni żagħżugħa (Fabri, dalwaqt jiġi ppubblikat[b]).

Fir-rigward tal-ammont ta' Malti fuq l-internet b'mod ġenerali, huwa diffiċli biex toħroġ b'numri eżatti, mhux l-anqas minħabba l-għadd ta' websajts qed jinbidel kontinwament. Iżda hemm fatturi oħra li jagħtu idea dwar l-ammont ta' Malti fuq l-internet meta mqabbel ma' lingwi oħra.

L-ewwel ħarsa lejn l-ammont ta' daħliet fil-Wikipedia (fl-1 ta' Ġunju, 2011) wera li kien hemm madwar 2,820 daħla bil-Malti

b'kuntrast ma' aktar minn 3,640,000 daħla bl-Ingliż u aktar minn 1,238,000 daħla bil-Ġermaniż.

Meta wieħed iqabbel in-numru tad-Dominju tal-Ogħla Livell (TLD), it-TLD .mt jokkupa l-pożizzjoni 213 (minn 358) b'numru mhux speċifikat ta' dominji .mt rreġistrati (membru taċ-Ċentru Informazzjoni tan-Network ta' Malta ta stima ta' madwar 5,000), imqabbel ma' 21,336,063 dominju irreġistrat għal .com (kummerċjali, klassifikazzjoni 1) u 5,459,604 dominji għad- .de (il-Ġermanja, klassifikazzjoni 2). Naturalment, in-numru ta' dominji rreġistrati ma jgħid xejn dwar il-lingwa li l-paġni taħt ċertu dominju huma miktubin biha.

Xi numri approssimattivi tal-ammont tal-lingwa Maltija fuq l-internet jistgħu jiġu kkalkulati permezz ta' proċedura proposta minn Kilgarriff u Grefenstette (2003) [xvi]. L-idea bażika hija li kliem funzjonali (eż. iżda, għal, dan eċċ.) huma aktar frekwenti minn kliem ta' kontenut (eż. nomi, verbi, aġġettivi) u jiffurma sett finit fil-lingwa. Barra minn hekk, il-persentaġġ ta' kliem funzjonali f'lingwa jkun stabbli f'kampjun ta' test meta d-daqs tal-kampjun jiżdied (il-Liġi Zipf). Għaldaqstant, wieħed jista' jikkalkula l-ammont ta' kliem għal kull lingwa fuq l-internet kif ġej:

L-ewwel nett, wieħed jikkalkula l-ammont ta' kliem funzjonali magħżula bil-Malti f'korpus (jiġifieri ġabra ta' testi) li d-daqs ikun magħruf. It-tieni nett, wieħed juża magna ta' tiftix (eż. Google) biex isib il-frekwenza għall-istess kliem funzjonali fuq il-web. Fit-tielet pass, il-frekwenza min-numru tal-korpus tiġi estrapolata għall-Google Search u mbagħad medja tiġi kkalkolata għall-frekwenza tal-kliem funzjonali fir-rizultati ta' tiftix.

Xi restrizzjonijiet ta' dan il-metodu għandhom jissemmew: L-ewwel nett, in-numri miksuba b'dan il-metodu huma biss paġni mtellgħin. Pereżempju, 94,300 paġna tal-Google għall-kelma *għal* 'for' mhumiex 94,300 każ tal-kelma fuq l-internet, iżda 94,300 paġni fuq il-web li fihom il-kelma *għal* mill-anqas darba. It-tieni, it-tfittxija ssib biss paġni fuq il-web li għandhom URL individwali (Kilgarriff u Grefenstette, 2003: p. 339). Paġni li huma aċċessibbli biss permezz ta' interface tal-web mhumiex miksuba bit-tiftix bl-internet. It-tielet, magna ta' tiftix tfittex biss għal sensiela ta' ittri irrispettivament mill-ambjent tagħha fuq il-paġna tal-web. Din ma tagħmel l-ebda ġudizzju dwar jekk ċerta sensiela ta' ittri hijiex tabilħaqq kelma ta' lingwa.

Il-metodu deskritt aktar kmieni, applikat għall-kliem funzjonali Maltin, jiġġenera stimi differenti għall-Malti. Għas-websajts mad-dominju .mt li jinsabu f'Malta, id-daqs stmat huwa ta' 50 miljun kelma, filwaqt li għas-websajts bid-dominju .mt fir-reġjuni kollha id-daqs huwa 500 miljun kelma. Ir-raġuni għal din id-differenza hija li ħafna dominji .mt huma riżervati għal servers barra minn Malta.

Ir-riżultati eżatti tat-tfittxijiet fil-Google (imwettqa fit-8 ta' Lulju, 2011) u l-estrapolazzjoni tagħhom tista' tiġi ttraċċata lura fit-Tabella A u t-Tabella B hawn taħt. Il-kolonna f/m (jiġifieri frekwenza għal kull miljun) tidentifika kemm il-kelma rispettiva sseħħ ta' spiss f'miljun, fil-korpus MLRS. Pereżempju, fit-Tabella A, il-kelma *għal* 'for' tidher kważi 3731 darba f'miljun kelma. It-tiftix tal-Google għall-kelma *għal* tirriżulta f'94,300 paġna b'mill-anqas okkażjoni waħda ta' *għal* fuq paġna web taħt id-dominju .mt

f'Malta. Multiplikazzjoni ta' miljun u diviżjoni b'3730.96 jagħti ammont stmat ta' 25,274,996 każ ta' kull kelma Maltija fuq il-paġni taħt id-dominju .mt ġewwa Malta. Jekk wieħed jagħmel dan il-kalkolu għall-kliem l-ieħor fit-tabella u jsib il-medja tar-riżultati, wieħed jasal għal numru ftit anqas minn 50 miljun kelma. Għall-paġni web madwar id-dinja kollha elenkati taħt id-dominju .mt, ir-riżultati huma għaxar darbiet ogħla.

Tabella A: Tfittxija bil-Google, ristretta għad-dominju .mt u r-reġjun ta' Malta

| Kelma | f/m | Google (.mt biss, Reġjun=mt) | Estrapolazzjoni |
|---|---|---|---|
| għal | 3730.96 | 94,300 | 25,274,996 |
| qed | 4770.79 | 118,000 | 24,733,849 |
| minn | 4833.58 | 173,000 | 35,791,276 |
| kien | 4073.83 | 93,800 | 23,025,015 |
| biex | 5276.78 | 179,000 | 33,922,202 |
| dan | 6412.28 | 434,000 | 67,682,634 |
| kienet | 1452.42 | 116,000 | 79,866,705 |
| kienu | 1465.56 | 135,000 | 92,114,959 |
| kont | 521.43 | 34,200 | 65,588,861 |
| konna | 301.39 | 19,400 | 64,368,426 |
| jekk | 2776.8 | 72,100 | 25,965,140 |
| mhux | 2101.32 | 79,500 | 37,833,362 |
| Medja | | | 48,013,952 |

Tabella B: Tfittxija bil-Google, ristretta għad-dominju .mt biss

| Kelma | f/m | Google (dominji kollha, Reġjun=mt) | Estrapolazzjoni |
|---|---|---|---|
| għal | 3730.96 | 1,340,000 | 359,156,892 |
| qed | 4770.79 | 966,00 | 202,482,188 |
| minn | 4833.58 | 1,240,000 | 256,538,632 |
| kien | 4073.83 | 3,100,000 | 760,954,679 |
| biex | 5276.78 | 6,530,000 | 1,237,497,110 |
| dan | 6412.28 | 3,980,000 | 620,684,062 |
| kienet | 1452.42 | 665,000 | 457,856,543 |
| kienu | 1465.56 | 436,000 | 297,497,202 |
| kont | 521.43 | 450,000 | 863,011,334 |
| konna | 301.39 | 81,600 | 270,745,546 |
| jekk | 2776.8 | 1,120,000 | 403,341,976 |
| mhux | 2101.32 | 1,040,000 | 494,926,998 |
| Medja | | | 518,724,430 |

Naturalment, għal studju serju, din it-tfittxija u l-estrapolazzjoni jkollhom jinkludu aktar kliem biex jaslu għal numri aktar affidabbli għall-ammont ta' Malti fuq l-internet. Iżda meta wieħed iqabbel ir-riżultati mat-Tabella 3 fuq p. 339 f'Kilgarriff & Grefenstette (2003), wieħed jista' jgħid li ż-żewġ numri huma baxxi ħafna: għall-paġni web f'Malta biss, l-ammont huwa aktar mil-Latvjan u anqas mill-Iżlandiż għaxar snin ilu (in-numri fit-Tabella 3 ġew ikkalkulati f'Marzu 2001). Għall-paġni web dinjin, l-ammont tal-Malti huwa aktar mill-Ungeriż u anqas miċ-Ċek għaxar snin ilu. Minħabba li "l-

proporzjon ta' testi mhux bl-Ingliż għall-Ingliż qed jikber" (Kilgarriff & Grefenstette, 2003: p. 339), il-Malti jista' jkun saħansitra rappreżentat anqas fuq l-internet illum mil-lingwi li għadhom kemm issemmew.

Apparti minn paġni ewlenin privati u weblogs, hemm numru ta' websajts uffiċjali bil-Malti. L-ewwel nett, hemm il-paġna ewlenija tal-gvern Malti[xvii], li hija disponibbli kemm bil-Malti kif ukoll bl-Ingliż. Barra minn hekk, hemm l-edizzjonijiet tal-internet tal-gazzetti ta' kuljum u ta' kull ġimgħa bil-lingwa Maltija: *In-Nazzjon, L-Orizzont* (kuljum), *Illum, Il-ĠENSillum, KullĦadd, Leħen is-Sewwa, It-Torċa* (kull ġimgħa).

Is-websajts tat-TV Malti u l-istazzjonijiet tar-radju juru taħlita tal-Ingliż u l-Malti fi gradi differenti. Pereżempju, is-websajts tal-istazzjonijiet tat-televiżjoni NET TV[xviii] u One TV[xix] għandhom qafas bl-Ingliż, flimkien ma' xi artikli bil-Malti, minkejja li l-programm tagħhom jinkludi titli kemm bil-Malti kif ukoll bl-Ingliż. L-istazzjon tar-radju tal-knisja RTK[xx] (Malti u Ingliż) jippermetti lill-utent jagħżel bejn iż-żewġ lingwi. Is-sitweb tal-Public Broadcasting Services (PBS)[xxi] fih sezzjonijiet bl-Ingliż u taqsimiet bil-Malti kif għandu s-sitweb tar-Radju 101[xxii]. Din it-taħlita bejn Ingliż u Malti tirrifletti l-użu tal-lingwa fil-ħajja ta' kuljum. Madankollu, fil-programmi, is-sitwazzjoni hija aktar ċara, minħabba li l-Awtorità tax-Xandir ta' Malta ħarġet linji gwida stretti għall-użu tal-Malti fuq it-TV u r-radju. Skont dawn, il-preżentaturi għandhom jitkellmu bil-Malti jew bl-Ingliż u mhux jaqilbu bejn iż-żewġ lingwi (Fabri, dalwaqt jiġi ppubblikat[a]: p. 28). Għaldaqstant il-programmi tal-istazzjonijiet jinkludu xandiriet bil-Malti biss u oħrajn bl-Ingliż biss. Dawn ikunu wkoll spiss disponibbli fuq l-internet, jew bħala live stream inkella podcasts.

Barra minn Malta, kollezzjoni kbira ta' testi bil-Malti tinsab fi ħdan il-EUR-Lex[xxiii] li tospita l-liġi uffiċjali u dokumenti oħra tal-Unjoni Ewropea mill-1951 fit-23 lingwa uffiċjali tagħha.

Ħafna jekk mhux id-dokumenti kollha tal-web disponibbli b'mod miftuħ jintużaw fi proġetti tal-korpus, eż. il-*JRC-Acquis Multilingual Parallel Corpus*[xxiv], li huwa korpus parallel li fih testi sħaħ tal-Liġi tal-Unjoni Ewropea fi 22 lingwa. Korpus ieħor li fih numru dejjem jikber ta' dokumenti tal-web viżibbli bil-Malti huwa l-korpus tal-MLRS (Server għar-Riżorsi Lingwistiċi bil-Malti) [xxv].

## Referenzi


Ambros, Arne (1998): *Bonġornu, kif int? Einführung in die maltesische Sprache*. Wiesbaden: Reichert.

Aquilina, Joseph (1987): *Maltese-English Dictionary Vol. I, A-L*. Malta: Midsea Books.

Aquilina, Joseph (1990): *Maltese-English Dictionary Vol. II, M-Z*. Malta: Midsea Books.

Biber, Douglas (1991): *Variation across speech and writing*. Cambridge: Cambridge University Press.

Borg, Albert J. & Azzopardi-Alexander, Marie (1997): *Maltese*. Routledge.

Brincat, Joseph M. (1995): *Malta 870-1054: Al-Himyarī's Account and its Linguistic Implications*. Malta: Said International.

Brincat, Joseph M. (2011): *Maltese and other Languages: a Linguistic History of Malta*. Malta: Midsea Books.

Bovingdon, Roderick (2001): *The Maltese language of Australia: Maltraljan: a lexical compilation with linguistic notations & a social, political and historical background*. Munich: LINCOM Europa.

Camilleri, Antoinette (1995): *Bilingualism in Education: The Maltese Experience*. New York: Brill.

Ebert, Karen (2000): "Aspect in Maltese." In: Dahl, Östen (ed.): *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 753-788.

Fabri, Ray (1993): *Kongruenz und die Grammatik des Maltesischen*. Tübingen: Niemeyer.

Fabri, Ray (1995): "The Tense and Aspect System of Maltese." In: R. Thieroff (ed.): *Tempussysteme in europäischen Sprachen II*. Niemeyer, Tübingen, 327-343.

Fabri, Ray (dalwaqt jiġi ppubblikat[a]): "Maltese" In: C. Delcourt and P. van Sterkenburg (eds.) The Languages of the 25. Revue belge de Philologie et d'Histoire: RBPH, 85 (2007) 3 and Amsterdam-Philadelphia: John Benjamins.

Fabri, Ray (dalwaqt jiġi ppubblikat[b]): "The Language of Young People and Language Change in Maltese." *In: Variation and Change: The dynamics of Maltese in space, time, and society* edited by Sandro Caruana, Ray Fabri & Thomas Stolz. Berlin: Akademie Verlag, 2011.

Kilgarriff, A. and Grefenstette, G. (2003): "Introduction to the special issue on the web as corpus", Computational Linguistics, 29: 333-347.

Kontzi, Reinhold (2005): *Sprachkontakt im Mittelmeer: Gesammelte Aufsätze zum Maltesischen*. Narr, Tübingen.

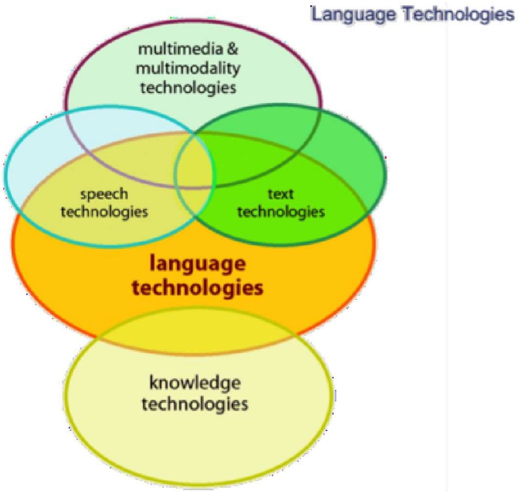
Il-Kunsill Nazzjonali tal-Ilsien Malti (2008a): *Deċiżjonijiet 1 tal-Kunsill Nazzjonali tal-Ilsien Malti dwar il-Varjanti Ortografiċi*. Il-Furjana.

Il-Kunsill Nazzjonali tal-Ilsien Malti (2008b): *Innaqqsu l-Inċertezzi*.

Mifsud, M. (1995): *Loan verbs in Maltese: a descriptive and comparative study*. Leiden etc: Brill.

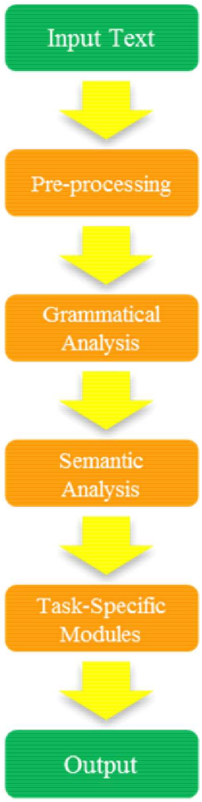
# Appoġġ ta' Teknoloġija Lingwistika għall-Malti

## Teknoloġiji Lingwistiċi




Teknoloġiji lingwistiċi huma teknoloġiji ta' informazzjoni li huma speċjalizzati biex jittrattaw il-lingwa umana. Għalhekk dawn it-teknoloġiji huma wkoll ta' spiss ikorporati taħt it-terminu Teknoloġija Lingwistika Umana. Il-lingwa umana sseħħ fil-forma mitkellma u miktuba. Filwaqt li t-taħdit huwa l-eqdem u l-aktar mod naturali ta' komunikazzjoni lingwistika, l-informazzjoni kumplessa u l-biċċa l-kbira tal-għarfien uman jinżammu u jiġu trażmessi f'testi miktubin. Teknoloġiji ta' taħdit u testi jipproċessaw jew jipproduċu lingwa f'dawn iż-żewġ modi ta' produzzjoni. Iżda l-lingwa għandha wkoll aspetti li huma maqsuma bejn taħdit u testi bħal dizzjunarji, il-biċċa l-kbira tal-grammatika u t-tifsir ta' sentenzi. Għaldaqstant partijiet kbar mit-teknoloġija lingwistika ma jistgħux jiġu inklużi taħt teknoloġiji tat-taħdit inkella tat-testi. Fost dawn hemm teknoloġiji li jorbtu l-lingwa mal-għarfien. Ir-rappreżentazzjoni fuq il-lemin turi l-pajsaġġ grafiku tat-Teknoloġija Lingwistika. Fil-komunikazzjoni tagħna aħna nħalltu l-lingwa ma' modi oħra ta' komunikazzjoni u mezzi oħra ta' informazzjoni. Aħna norbtu t-taħdit ma' ġesti u espressjonijiet tal-wiċċ. Testi diġitali jingħaqdu ma' stampi u ħsejjes. Il-films jistgħu jinkludu lingwa f'forma mitkellma u miktuba. Għaldaqstant teknoloġiji ta' taħdit u testi jikkoinċidu u jinteraġixxu ma' ħafna teknoloġiji oħra li jiffaċilitaw l-ipproċessar ta' komunikazzjoni multimodali u dokumenti multimidjali.

## Arkitetturi ta' Applikazzjonijiet għat-Teknoloġija Lingwistika




Applikazzjonijiet ta' software tipiċi għall-ipproċessar tal-lingwa jikkonsistu f'diversi komponenti li jirriflettu aspetti differenti tal-lingwa u tal-kompitu li jkunu qed jimplimentaw. Ir-rappreżentazzjoni fuq il-lemin turi arkitettura ferm simplifikata li tista' tinsab f'sistema ta' pproċessar ta' testi. L-ewwel tliet moduli jittrattaw l-istruttura u t-tifsira tat-test imdaħħal fis-sistema:

- Qabel l-ipproċessar: tindif tad-dejta, tneħħija tal-ifformattjar fejn hu xieraq, sejbien tal-lingwa mdaħħla fis-sistema, standardizzazjoni tar-rappreżentazzjoni ta' simboli speċjali bħas-sing fil-Malti
- Analiżi grammatikali:  tiftix tal-verb u l-oġġetti tiegħu, il-modifikaturi, eċċ;  sejbien tal-istruttura tas-sentenza.
- Analiżi semantika: tneħħja tal-ambigwità (Liema tifsira ta' "bank" hija t-tajba f'kuntest partikolari?), soluzzjoni ta' anafori u espressjonijiet ta' referenza bħal "hi", "il-karozza", eċċ; rappreżentazzjoni tat-tifsira tas-sentenza b'mod li tinqara minn magna

Moduli ta' kompiti speċifiċi mbagħad iwettqu ħafna operazzjonijiet differenti bħal sommarju awtomatiku ta' test imdaħħal fis-sistema, tfittxija ta' bażi ta' dejta u ħafna oħrajn. Hawn taħt, se nagħtu eżempji ta' **oqsma ewlenin ta' applikazzjoni** u niġbdu l-attenzjoni għal xi wħud mill-moduli ta' arkitetturi differenti f'kull sezzjoni. Għal darb'oħra, l-arkitetturi huma simplifikati ferm u

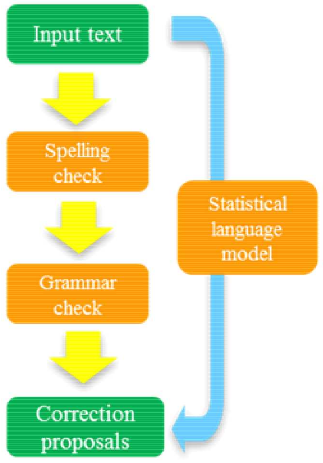
idealizzati, sabiex iservu biex wieħed jispjega l-kumplessità tal-applikazzjonijiet tat-teknoloġija lingwistika b'mod ġenerali li jinftiehem.

Wara l-introduzzjoni tal-oqsma ewlenin ta' applikazzjoni, se nagħtu ħarsa ġenerali fil-qosor lejn is-sitwazzjoni fir-riċerka u l-edukazzjoni tat-TL, u nikkonkludu b'deskrizzjoni ta' programmi ta' ffinanzjar (tal-passat). Fl-aħħar ta' din it-taqsima, se nippreżentaw stima esperta dwar is-sitwazzjoni fir-rigward ta' għodod prinċipali u riżorsi tat-TL f'numru ta' dimensjonijiet bħal disponibbiltà, maturità, jew kwalità. Din it-tabella tagħti ħarsa ġenerali tajba lejn is-sitwazzjoni tat-TL għall-Malti.

## Oqsma ewlenin ta' applikazzjoni

### L-Iċċekkjar tal-lingwa

Kull min juża għodda ta' pproċessar tal-kliem bħal Microsoft Word iltaqa' ma' komponent li jiċċekkja l-ortografija, jindika żbalji ortografiċi u jipproponi korrezzjonijiet. 40 sena wara l-ewwel programm ta' korrezzjoni tal-ortografija minn Ralph Gorin, ċekkjaturi lingwistiċi llum ma jqabblux biss il-lista ta' kliem estratta ma' dizzjunarju ta' kliem spellut b'mod korrett, iżda saru dejjem aktar sofistikati. Minbarra l-algoritmi li jiddependu fuq il-lingwa għall-immaniġġjar tal-morfoloġija (eż. formazzjoni tal-plural), xi wħud issa huma kapaċi jagħrfu żbalji marbuta ma' sintassi, bħal verb nieqes jew verb li ma jaqbilx mas-suġġett tiegħu fil-persuna u n-numru, eż. f' *'She *write a letter.'* Madankollu, iċ-ċekkjaturi l-aktar disponibbli (inkluż Microsoft Word) ma jsibu l-ebda żbalji fl-ewwel vers ta' poeżija minn Jerrold H. Zar (1992) li tidher hawn taħt:




*Eye have a spelling chequer,*

*It came with my Pea Sea.*

*It plane lee marks four my revue*

*Miss Steaks I can knot sea.*

Għall-immaniġġjar ta' dan it-tip ta' żbalji, l-analiżi tal-kuntest huwa meħtieġ f'ħafna każijiet, eż. sabiex jiġi deċiż f'liema pożizzjoni l-għ siekta għandha tinkiteb f'verb Malti, bħal f'dan l-eżempju:

a) *...in-negozjati li kien għamel il-Gvern ...*

b) *Pawlu, agħmel l-eżamijiet!*

c) **...in-negozjati li kien agħmel il-Gvern ...*

Iż-żewġ verbi *għamel* u *agħmel* jiġu ppronunzjati [ɐː.mɛl].

Dan jew jeħtieġ il-formulazzjoni ta' regoli tal-grammatika li jkunu speċifiċi, jiġifieri livell għoli ta' kompetenza u xogħol manwali, jew l-użu tal-hekk imsejjaħ mudell lingwistiku tal-istatistika. Mudelli

bħal dawn jikkalkulaw il-probabbiltà ta’ kelma partikolari li tidher f’ambjent speċifiku (jiġifieri, il-kliem ta’ qabel u ta’ wara). Pereżempju, *kien għamel* hija sekwenza ta’ kliem ħafna aktar probabbli milli *kien agħmel*. Mudell lingwistiku tal-statistika jista’ jitnissel awtomatikament minn ammont kbir ta’ dejta lingwistika (xierqa) (jiġifieri korpus). Sa issa, dawn il-metodi ġew l-aktar żviluppati u evalwati fuq dejta lingwistika bl-Ingliż. Madankollu, dawn mhux neċessarjament jittrasferixxu sew għal-lingwi li jkollhom inflessjoni għolja bħall-Malti, fejn tip ta’ kelma partikolari, bħal verb, tista’ tagħti numru kbir ta’ forom ortografiċi.

L-użu ta’ Ċekkjaturi Lingwistiċi mhuwiex limitat għal għodod ta’ pproċessar tal-kliem, iżda jiġi wkoll applikat f’sistemi ta’ appoġġ għal min jikteb. L-ammont ta’ dokumentazzjoni teknika li jakkumpanja l-għadd dejjem jikber ta’ prodotti tekniċi, żdied b’mod mgħaġġel matul l-aħħar deċennji. Minħabba l-biża’ tal-ilmenti tal-klijenti dwar l-użu ħażin u talbiet għal danni li jirriżultaw minn istruzzjonijiet ħżiena jew mifhuma ħażin, il-kumpaniji bdew jiffokaw dejjem aktar fuq il-kwalità tad-dokumentazzjoni teknika, fl-istess ħin jimmiraw għas-suq internazzjonali. Avvanzi fl-ipproċessar lingwistiku naturali wasslu għall-iżvilupp ta’ software ta’ appoġġ għal min jikteb, li jgħin lill-kittieb ta’ dokumentazzjoni teknika sabiex juża vokabolarju u strutturi ta’ sentenzi konsistenti ma’ ċerti regoli u restrizzjonijiet (korporattivi) tat-terminoloġija.

Bħal lingwi oħra, mezz biex jiddetermina jekk sensiela partikolari hijiex kelma valida mhux kundizzjoni suffiċjenti għal sejbien ta’ żbalji ortografiċi, iżda huwa kundizzjoni neċessarja. S’issa, għalkemm saru diversi tentattivi, l-ebda mezz bħal dan ma jeżisti għall-Malti.

Wieħed minn tal-ewwel kien ta’ Mangion (1999) li juża forom rudimentali ta’ analiżi morfoloġiċi bbażati fuq regoli. Kelma kienet essenzjalment ikkunsidrata valida jekk tistax tkun derivata permezz ta' zokk misjub f’dizzjunarju. Il-problema b’ dan il-metodu huwa li jeħtieġ lista kompluta ta’ kull zokk, u naturalment, ir-regoli għandhom ikunu preċiżi ħafna. Ir-riżultati kienu kemxejn limitati mil-lista ta’ zkuk, li ma kinitx kompluta, u n-natura imperfetta tar-regoli.

Metodu ieħor ħares lejn l-istatistiċi għal soluzzjoni. L-idea intuwittiva hija li għal lingwa partikolari, ċerti sekwenzi ta’ karattri huma improbabbli ħafna. Bl-Ingliż, pereżempju, qatt ma nsibu s-sekwenza “kk”, għalhekk jekk isseħħ taħt l-istess sekwenza f’kelma miktuba, nistgħu nbassru, bi grad għoli ta’ kunfidenza, li l-kelma mhijiex valida. B’mod aktar ġenerali, nistgħu nikkalkulaw il-probabbiltà ta’ xi sekwenza bħala funzjoni tal-probabbiltajiet tas-sekwenzi ta’ taħtha kollha, bl-adozzjoni tal-prinċipju li sabiex il-kelma tiġi kkunsidrata valida, il-probabbiltà trid taqbeż ċertu limitu.

Ċekkjatur ortografiku statistiku li jagħmel użu minn prinċipju bħal dan kien żviluppat minn Mizzi (2001). Dan ma kienx jeħtieġ dizzjunarju, iżda minflok kien ibbażat fuq id-distribuzzjoni ta’ n-grammi ta' karattri misjuba f’korpus ta’ gazzetta. Deher ċar li biex dan il-metodu jirnexxi kien jeħtieġ (i) mudell lingwistiku aktar preċiż li jirrikjedi aktar dejta lingwistika minn dik disponibbli dak iż-żmien, u (ii) li l-probabbiltà tas-sekwenza waħedha ma kinitx biżżejjed sabiex kelma ortografika tiġi kklassifikata bħala żball. Kif

issuġġerit hawn fuq, informazzjoni oħra hija meħtieġa, bħal kategorija tal-kliem mill-kuntest tal-madwar.

Tentattivi oħra biex jiġi żviluppat ċekkjatur ortografiku għall-Malti jinkludu ċekkjatur fuq l-internet li ġie żviluppat minn Ramon Casha tal-Linux User Group [xxvi]. Dan huwa bbażat fuq lista ta' kliem b'madwar miljun tip ta' kelma oriġinarjament miġbura minn korpus li jvarja, u sussegwentement estiż permezz ta' regoli differenti għall-immaniġġjar ta' inflessjonijiet. L-eżattezza tiegħu ma ġietx stabbilita uffiċjalment. Microsoft ukoll kienu qed jaħdmu fuq ċekkjatur ortografiku biex jinkluduh mal-pakkett ta' interface tagħhom għal-lingwa Maltija għalkemm mhux magħruf meta dan se jiġi rilaxxat.

L-użu ta' ċekkjatur lingwistiku mhuwiex limitat għal għodod ta' pproċessar tal-kliem. Oqsma ta' applikazzjoni oħra jinkludu **appoġġ għal min jikteb**, pereżempju biex jgħin lill-kittieb ta' dokumentazzjoni teknika juża vokabolarju tekniku b'mod konsistenti, u l-qasam tat-**tagħlim tal-lingwi assistit mill-kompjuter**. L-iċċekkjar tal-lingwa jiġi wkoll applikat biex jiġu kkoreġuti awtomatikament mistoqsijiet mibgħuta lil magni ta' tiftix, eż. suġġerimenti tat-tip "Ridt tfisser ..." lil Google.

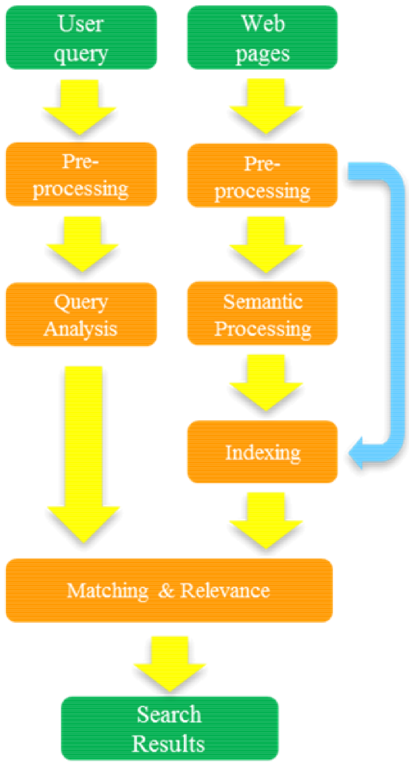
Barra minn CD ta' dizzjunarju interattiv bl-istampi (Sciriha 1997), sal-lum l-ebda applikazzjoni bħal din ma ġiet żviluppata għall-Malti.

### Tiftix fuq il-web

Tiftix fuq il-web, f'intranets, jew libreriji diġitali huwa probabbilment l-aktar teknoloġija lingwistika użata llum, iżda l-anqas żviluppata. Il-magna ta' tiftix Google, li bdiet fl-1998, illum hija użata għal madwar 80% tat-tiftix kollu fid-dinja kollha[xxvii]. Mill-2004, il-verb google huwa saħansitra msemmi fid-dizzjunarju *Cambridge Advanced Learner's Dictionary*. La l-interface tat-tiftix u lanqas il-preżentazzjoni tar-riżultati miksuba ma nbidlu b'mod sinifikanti mill-ewwel verżjoni. Fil-verżjoni kurrenti, Google tipproponi korrezzjoni tal-ortografija għall-kliem spellut ħażin u, fl-2009, inkorporat kapaċitajiet bażiċi ta' tiftix semantiku fit-taħlita algoritmika tagħha[xxviii], li tista' ttejjeb il-preċiżjoni tat-tiftix billi tanalizza t-tifsira tat-termini mistoqsija f'kuntest. Is-suċċess tal-istorja tal-Google turi li b'ammont sostanzjali ta' dejta disponibbli u tekniki effiċjenti għall-indiċjar ta' din id-dejta, metodu bbażat il-biċċa l-kbira fuq statistika, jista' jwassal għal riżultati sodisfaċenti.

Madankollu, għal talba aktar sofistikata għall-informazzjoni, l-integrazzjoni ta' għarfien aktar profond tal-lingwistika hija essenzjali. Fil-laboratorji ta' riċerka, esperimenti bl-użu ta' teżawri li jinqraw minn magni u riżorsi ontoloġiċi tal-lingwa bħal WordNet urew titjib billi jippermettu s-sejbien ta' paġna fuq il-bażi ta' sinonimi tat-termini ta' tiftix, eż. *Atomkraft, Kernenergie u Nuklearenergie* (enerġija atomika, enerġija atomika/nukleari, u enerġija nukleari) jew saħansitra termini relatati aktar mill-bogħod.

Il-ġenerazzjoni li jmiss ta' magni ta' tiftix se jkollhom jinkludu teknoloġija lingwistika ferm aktar sofistikata. Jekk mistoqsija ta' tiftix tikkonsisti f'domanda jew tip ieħor ta' sentenza minflok lista ta' kliem ewlieni, il-ksib ta' tweġibiet rilevanti għal din il-mistoqsija teħtieġ analiżi ta' din is-sentenza fuq livell sintattiku u semantiku



kif ukoll id-disponibiltà ta' indiċi li jippermetti rkupru mgħaġġel tad-dokumenti rilevanti. Pereżempju, immaġina utent idaħħal il-mistoqsija "Tini lista tal-kumpaniji kollha li ttieħdu minn kumpaniji oħra fl-aħħar ħames snin". Għal tweġiba sodisfaċenti, parsing sintattika jeħtieġ li tiġi applikata biex tiġi analizzata l-istruttura grammatikali tas-sentenza u jiġi determinat li l-utent qed ifittex kumpaniji li ttieħdu minn kumpaniji oħra. Barra minn hekk, l-espressjoni *l-aħħar ħames snin* jeħtieġ li tiġi pproċessata biex jinstab liema snin qed tirreferi għalihom.

Fl-aħħar nett, il-mistoqsija pproċessata teħtieġ li tiġi mqabbla ma' ammont enormi ta' dejta mhux strutturata sabiex jinstabu l-biċċa jew biċċiet ta' informazzjoni li l-utent qed ifittex. Dan huwa ġeneralment imsejjaħ rkupru ta' informazzjoni (RI) u jinvolvi t-tfittxija għal u l-klassifikazzjoni ta' dokumenti rilevanti. Barra minn hekk, meta niġġeneraw lista ta' kumpaniji, jenħtieġ li niġbru l-informazzjoni ta' sekwenza partikolari ta' kliem f'dokument li tirreferi għal isem ta' kumpanija. Din it-tip ta' informazzjoni tkun disponibbli mill-hekk imsejħa identifikazzjoni ta' entità bl-isem.

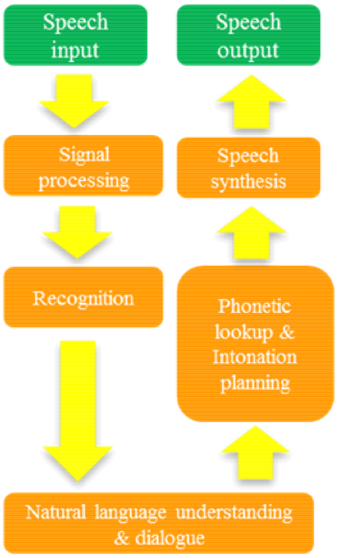
Saħansitra aktar diffiċli huwa t-tentattiv biex inqabblu mistoqsija ma' dokumenti miktuba f'lingwa differenti. Għall-irkupru ta' informazzjoni bejn il-lingwi, għandna nittraduċu awtomatikament il-mistoqsija fil-lingwi sorsi kollha possibbli u nittrasferixxu l-informazzjoni miksuba lura għal-lingwa fil-mira. Il-perċentwal dejjem jiżdied ta' dejta disponibbli f'formati mhux testwali imexxi d-domanda għal servizzi li jippermettu rkupru ta' informazzjoni multimidjali, jiġifieri, it-tiftix ta' informazzjoni li jinkludu immaġini, audio u dejta bil-videos. Għal files audio u video, dan jinvolvi modulu ta' identifikazzjoni ta' taħdit biex ibiddel il-kontenut tat-taħdit f'test jew rappreżentazzjoni fonetika, li magħhom jistgħu jitqabblu l-mistoqsijiet tal-utent.

F'Malta, hemm numru ta' websajts ta' tiftix li huma speċifikament immirati għal Malta[xxix]. Barra minn hekk, hemm numru żgħir ta' SMEs ibbażati f'Malta li jinkorporaw tekniki relattivament sofistikati tal-ipproċessar tal-lingwa fl-ambitu ta' applikazzjonijiet ta' tiftix. Charonite[xxx], pereżempju, hija SME lokali li tittratta l-aħjar użu ta' magni ta' tiftix. Madankollu, bħalissa m'hemm l-ebda magni ta' tiftix kummerċjalment disponibbli li huma speċifikament immirati lejn l-ilsien Malti, apparti minn prototip għall-irkupru ta' informazzjoni bejn il-lingwi żviluppat għal skopijiet tal-LT4eL[xxxi], proġett ta' riċerka Ewropew tal-FP6 li uża għodod ta' teknoloġija lingwistika multilingwi u tekniki ta' kodifikazzjoni ta' semantika għal titjib fil-ksib ta' materjal għat-tagħlim.

### L-Interazzjoni tat-taħdit

Teknoloġija għall-interazzjoni tat-taħdit hija l-bażi għall-ħolqien ta' interfaces li jippermettu lill-utent biex jinteraġixxi ma' magni billi juża l-lingwa mitkellma aktar milli, pereżempju, stampa grafika, tastiera, u mouse. Illum, l-interfaces għall-vuċi (VUIs) huma normalment użati għal offerti ta' servizzi ta' awtomatizzazzjoni parzjali jew kompluti pprovduti mill-kumpaniji lill-klijenti tagħhom, l-impjegati jew l-imsieħba permezz tat-telefon. Dominji ta' negozju li jiddependu ħafna fuq VUIs huma l-banek, il-loġistika, it-trasport pubbliku, u t-telekomunikazzjonijiet. Użi oħra tat-Teknoloġija għall-Interazzjoni tat-Taħdit huma interfaces għal apparat partikolari, eż. sistemi ta' navigazzjoni fil-karozza, u l-użu

tal-lingwa mitkellma bħala alternattiva għall-modalitajiet ta' input/output ta' interfaces grafiċi għall-utent, eż. fl-ismartphones.




Fil-qalba tagħha, l-Interazzjoni tat-Taħdit tinkludi l-erba' teknoloġiji differenti li ġejjin:

- L-Identifikazzjoni awtomatika tat-taħdit (ASR) hija responsabbli biex tiddetermina liema kliem kien attwalment mitkellem f'sekwenza partikolari ta' ħsejjes imlissna minn utent.
- L-analiżi sintattika u l-interpretazzjoni semantika janalizzaw l-istruttura sintattika tat-tlissin tal-utent u jinterpretaw tal-aħħar skont l-għan tas-sistema rispettiva.
- Il-ġestjoni tad-djalogu hija meħtieġa sabiex jiġi determinata ma' liema parti tas-sistema l-utent jinteraġixxi, liema azzjoni għandha tittieħed skont l-input tal-utent u l-funzjonalità tas-sistema.
- It-teknoloġija tas-sinteżi tat-taħdit (Text-to-Speech, TTS) hija użata biex tittrasforma l-kliem ta' dik l-espressjoni fi ħsejjes li jkunu prodotti għall-utent.

Waħda mill-isfidi ewlenin hija li jkollok sistema ASR li tidentifika l-kliem imlissen minn utent bl-aktar mod preċiż possibbli. Dan jeħtieġ jew restrizzjoni tal-firxa ta' espressjonijiet possibbli tal-utent għal sett limitat ta' kliem ewlieni, inkella l-ħolqien manwali ta' mudelli lingwistiċi li jkopru firxa kbira ta' espressjonijiet ta' lingwa naturali tal-utent. Filwaqt li tal-ewwel jirriżulta f'użu pjuttost riġidu u inflessibbli ta' VUI u li possibilment jikkawża aċċettazzjoni baxxa mill-utenti, il-ħolqien, l-irfinar u l-manutenzjoni tal-mudelli lingwistiċi jistgħu jżidu l-ispejjeż b'mod sinifikanti. Madankollu, il-VUIs li jużaw mudelli lingwistiċi u fil-bidu jippermettu lill-utent biex jesprimi l-intenzjoni tiegħu b'mod flessibbli – evokati, pereżempju, minn tislima bħal 'Kif nista' ngħinek' – għandha kemm rata ogħla ta' awtomazzjoni, kif ukoll aktar aċċettazzjoni mill-utenti u għalhekk jistgħu jitqiesu bħala vantaġġużi fuq il-metodu ta' djalogu dirett u anqas flessibbli. Għall-parti tal-output tal-VUI, il-kumpaniji għandhom tendenza li jużaw ħafna espressjonijiet irrekordjati minn qabel ta' kelliema professjonali – idealment korporattivi. Għal espressjonijiet statiċi, li fihom il-kliem ma jiddependix fuq il-kuntesti partikolari tal-użu jew id-dejta personali ta' utent partikolari, dan iwassal għal esperjenza rikka tal-utent. Madankollu, aktar ma l-espressjoni jkollha tikkunsidra kontenut dinamiku, aktar l-esperjenza tal-utent tista' tbati aktar minn prosodija fqira li tirriżulta minn files audio individwali marbutin ma' xulxin. B'kuntrast ma' dan, is-sistemi TTS tal-lum, jagħtu xhieda li huma superjuri, għalkemm jistgħu jintużaw aħjar, fir-rigward tan-naturalezza prosodika ta' espressjonijiet dinamiċi.

Rigward is-suq tat-teknoloġija għall-Interazzjoni tat-Taħdit, l-aħħar għaxar snin għaddew minn standardizzazzjoni qawwija tal-interfaces bejn il-komponenti ta' teknoloġiji differenti, kif ukoll bi standards għal ħolqien ta' prodotti ta' software partikolari għal ċerta applikazzjoni. F'dawn l-aħħar għaxar snin kien hemm ukoll konsolidazzjoni soda tas-suq, partikolarment fil-qasam tal-ASR u t-TTS. Hawnhekk, is-swieq nazzjonali fil-pajjiżi tal-G20 – jiġifieri pajjiżi ekonomikament b'saħħithom b'popolazzjoni konsiderevoli -

huma dominati minn anqas minn 5 parteċipanti dinjija, b'Nuance u Loquendo dawk l-aktar prominenti fl-Ewropa.

Ħafna mill-iżvilupp tat-teknoloġija tat-taħdit f'Malta kkonċentra fuq 'mit-test għat-taħdit' (TTS). Xi xogħol pijunier fil-bidu kien imwettaq minn P. Micallef (1997) u dan kien segwit minn numru ta' teżijiet f'livell ta' Masters (Farrugia, 2005). Xi xogħol fuq sistema TTS ibbażata fuq il-web beda minn Buhagiar u Micallef (2008).

Żvilupp sinifikanti fis-sinteżi tat-taħdit għall-Malti kien il-kisba ta' offerta mingħand il-gvern għall-iżvilupp ta' sintetizzatur tat-taħdit mill-kumpanija lokali Crimson Wing Ltd. Malta. Dan ix-xogħol huwa parzjalment iffinanzjat mill-Fond Ewropew għall-Iżvilupp Reġjonali u kkummissjonat mill-Fondazzjoni għall-Aċċessibilità għat-Teknoloġija tal-Informazzjoni (FITA). Il-prototip se jkun konformi ma' SAPI u se jinkludi tliet vuċijiet (tal-irġiel, tan-nisa, u tat-tfal). Skont preżentazzjoni riċenti (Borg et al. 2011) ix-xogħol qed javvanza sew u prototip, mistenni fl-2012, se jkun disponibbli biex jitniżżel mingħajr ħlas.

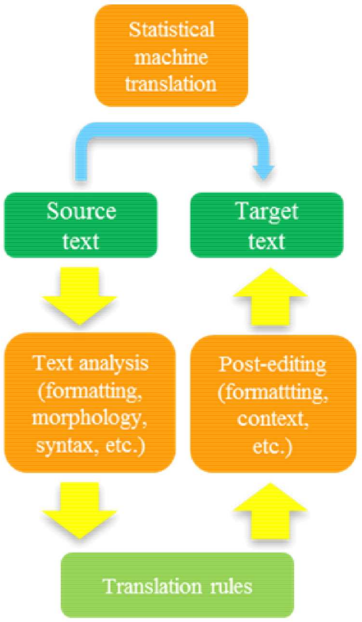
Ix-xogħol fuq l-idenifikazzjoni tat-taħdit huwa anqas avvanzat. Prototip biex jidentifika n-numri nħoloq minn (Calleja 2004) f'dominji sempliċi. Fir-rigward tat-taħdit, il-problema fundamentali tibqa' nuqqas ta' dejta annotata b'mod xieraq minħabba li dan jeħtieġ sforz manwali sinifikanti. Xi tentattivi ta' dħul awtomatiku saru minn Psaila (2008). Il-ħolqien ta' korpus u qafas deskrittiv għall-istudju tal-intonazzjoni Maltija bdiet mill-Istitut tal-Lingwistika u twettqet minn Vella u Farrugia (Vella u Farrugia 2006). Huwa mistenni li l-korpus li qed jiġi żviluppat minn Crimson Wing se jkun disponibbli għar-riċerka.

Ħarsa lil hinn mill-istat tat-teknoloġija tal-lum, turi li se jkun hemm bidliet sinifikanti minħabba l-firxa ta' smartphones bħala pjattaforma ġdida għall-ġestjoni ta' relazzjonijiet mal-klijenti – minbarra t-telefon, l-internet u mezzi għall-posta elettronika. Din it-tendenza se taffettwa wkoll l-użu tat-teknoloġija għall-Interazzjoni tat-Taħdit. Minn naħa, id-domanda għal telefonija bbażata fuq VUIs se tonqos, fuq medda ta' tul ta' żmien. Min-naħa l-oħra, l-użu tal-lingwa mitkellma bħala modalità ta' input faċli għall-utent għall-ismartphones se jikseb importanza sinifikanti. Din it-tendenza hija appoġġjata minn titjib osservabbli tal-eżattezza tal-identifikazzjoni tad-diskors li hija indipendenti mill-kelliem għal dettatura ta' taħdit li diġà hija offruta bħala servizzi ċentralizzati għall-utenti tal-ismartphone. B'din l-'esternalizzazzjoni' tal-kompitu ta' identifikazzjoni tal-infrastruttura tal-applikazzjonijiet, l-użu ta' applikazzjoni speċifika ta' teknoloġiji ewlenin lingwistiċi għandha tikber fl-importanza meta mqabbla mas-sitwazzjoni preżenti.

### Traduzzjoni awtomatika

L-idea tal-użu ta' kompjuters diġitali għat-traduzzjoni ta' lingwi naturali bdiet fl-1946 minn A. D. Booth u kienet segwita minn fondi sostanzjali għar-riċerka f'dan il-qasam fil-ħamsinijiet u ssoktat mill-ġdid fit-tmeninijiet. Madankollu, it-Traduzzjoni Awtomatika (TA) tibqa' tonqos li tissodisfa l-aspettativi għolja li ħolqot fis-snin bikrin tagħha.

Fil-livell bażiku tagħha, TA sempliċiment tissostitwixxi l-kliem ta' lingwa naturali waħda b'ta' oħra. Dan jista' jkun utli f'dominji ta' suġġetti b'lingwa ristretta u konvenzjonali ħafna, pereżempju, rapporti tat-temp. Madankollu, għal traduzzjoni tajba ta' testi anqas standardizzati, unitajiet ta' testi akbar (frażijiet, sentenzi, jew anki siltiet sħaħ) jeħtieġ li jkunu mqabbla mal-eqreb kontropartijiet tagħhom fil-lingwa fil-mira. Id-diffikultà prinċipali hawnhekk tinsab fil-fatt li l-lingwa umana hija ambigwa, u toħloq sfidi fuq livelli multipli, eż. tneħħija ta' ambigwità mis-sens tal-kelma fuq livell lessikali ('Jaguar' tista' tfisser karozza jew annimal) jew it-twaħħil ta' frażijiet prepożizzjonali fuq livell sintattiku bħal fl-eżempji li ġejjin:




*Il-Kuntistabbli osserva lir-raġel bit-teleskopju.*

*Il-Kuntistabbli osserva lir-raġel bir-rivolver.*

Mod wieħed kif wieħed jitratta l-kompitu huwa bbażat fuq regoli lingwistiċi. Għat-traduzzjonijiet bejn lingwi marbutin flimkien mill-qrib, traduzzjoni diretta tista' tkun possibbli f'każijiet bħall-eżempju t'hawn fuq. Iżda sistemi bbażati fuq regoli (jew immexxija minn għarfien) ta' spiss janalizzaw it-test imdaħħal fis-sistema u joħolqu rappreżentazzjoni intermedjarja u simbolika, li minnhom it-test fil-lingwa fil-mira jiġi ġġenerat. Is-suċċess ta' dawn il-metodi jiddipendi ħafna fuq id-disponibbiltà ta' dizzjunarji estensivi b' informazzjoni morfoloġika, sintattika, u semantika, u settijiet kbar ta' regoli tal-grammatika mfassla bir-reqqa minn lingwista tas-sengħa.

Lejn it-tmiem tat-tmeninijiet, kif l-enerġija kompjutazzjonali żdiedet u saret anqas għalja, kien hemm aktar interess fil-mudelli statistiċi għat-TA. Il-parametri ta' dawn il-mudelli statistiċi jittieħdu mill-analiżi ta' korpora ta' testijiet bilingwi, bħalma hu l-korpus parallel tal-Europarl, li fih il-proċedimenti tal-Parlament Ewropew fi 11-il lingwa Ewropea. B'dejta suffiċjenti, TA statistika jaħdem tajjeb biżżejjed li jikseb tifsira approssimattiva għal test ta' lingwa barranija. Madankollu, bid-differenza ta' sistemi mmexxija mill-għarfien, it-TA statistika (jew immexxija minn dejta) spiss tiġġenera produzzjoni mhux grammatikali. Min-naħa l-oħra, minbarra l-vantaġġ li anqas sforz uman huwa meħtieġ għall-kitba grammatikali, TA mmexxija minn dejta tista' wkoll tkopri partikolaritajiet tal-lingwa li jintilfu f'sistemi mmexxija minn għarfien, pereżempju espressjonijiet idjomatiċi.

Minħabba li l-punti tajbin u dgħajfin tat-TA mmexxija minn għarfien u dejta huma kumplimentari, ir-riċerkaturi llum unanimament jimmiraw għal metodi ibridi li jgħaqqdu l-metodoloġiji tat-tnejn li huma. Dan jista' jsir b'modi diversi. Wieħed jinkludi l-użu ta' kemm sistemi mmexxija minn għarfien kif ukoll dejta u jkollu modulu ta' għażla li jiddeċiedi dwar l-aħjar output għal kull sentenza. Madankollu, għal sentenzi itwal, l-ebda riżultat mhu se jkun perfett. Soluzzjoni aħjar hija li tgħaqqad l-aħjar partijiet ta' kull sentenza minn outputs multipli, li jistgħu jkunu pjuttost kumplessi, minħabba li partijiet korrispondenti ta' alternattivi multipli mhux dejjem huma ovvji u jridu jkunu allinjati.

F'Malta l-ħidma mwettqa fit-Traduzzjoni Awtomatika kienet ristretta għal ftit teżijiet tal-grad ta' Baċellerat u Masters. Sistema ta' trasferiment ibbażata fuq l-LFG kienet żviluppata għall-

Ingliż/Malti minn Farrugia (2000) u kienet tittraduċi b'suċċess it-tbassir tat-temp. Aktar tard J. Bajada (Bajada 2004, 2009) ħadem fuq TA statistika (SMT) bl-enfasi fuq tekniki għall-produzzjoni ta' mudelli lingwistiċi u tat-traduzzjoni. Ix-xogħol preċedenti kien jikkonċerna mudelli bbażati fuq kliem, filwaqt li tal-aħħar żviluppa tekniki għal ġbir ta' dejta bilingwali ta' frażijiet minn korpus limitat.

Bħal f'bosta oqsma oħra, il-problema bażika hija n-nuqqas ta' kwantitajiet kbar ta' dejta bilingwali annotata b'mod xieraq. Minħabba din ir-raġuni, forsi, is-sistema bil-punt ta' referenza biex jiġu ġġudikati avvanzi tibqa' Google Translate.

Il-kwalità tas-sistemi tat-TA hija meqjusa li għad għandha potenzjal kbir ta' titjib. L-isfidi jinkludu l-adattabilità tar-riżorsi lingwistiċi għal dominju ta' suġġett partikolari jew qasam tal-utent u l-integrazzjoni tax-xogħol għaddej eżistenti b'bażijiet ta' termini u memorji tat-traduzzjoni. Barra minn hekk, ħafna mis-sistemi kurrenti huma bbażati fuq l-Ingliż u jappoġġjaw lil ftit biss mil-lingwi minn u għall-Ġermaniż, li jwassal għal tensjonijiet fix-xogħol għaddej totali tat-traduzzjoni, u eż. jisforza utenti tat-TA biex jitgħallmu għodod ta' kodiċi ta' lessiku differenti għal sistemi differenti.

Kampanji ta' evalwazzjoni jippermettu tqabbil tal-kwalità tas-sistemi tat-TA, il-metodi diversi u l-istatus ta' sistemi TA għal-lingwi differenti. It-tabella 1, ippreżentata fi ħdan il-proġett Euromatrix+ tal-KE, turi l-prestazzjonijiet par par miksuba għal 22 lingwa uffiċjali (il-Gaelic Irlandiż huwa nieqes) f'termini ta' punteġġ BLEU[xxxii].

L-aħjar riżultati (murija bl-aħdar u l-blu) nkisbu minn lingwi li jibbenefikaw minn sforzi konsiderevoli ta' riċerka, fi ħdan programmi kkoordinati, u mill-eżistenza ta' korpora paralleli u numerużi (eż. Ingliż, Franċiż, Olandiż, Spanjol, Ġermaniż), l-agħar (bl-aħmar) minn lingwi li ma bbenefikawx minn sforzi simili, jew li huma differenti ħafna minn lingwi oħra (eż. Ungeriż, Malti, Finlandiż).

Target Language

| | en | bg | de | cs | da | el | es | et | fi | fr | hu | it | lt | lv | mt | nl | pl | pt | ro | sk | sl | sv |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| en | – | 40.5 | 46.8 | 52.6 | 50.0 | 41.0 | 55.2 | 34.8 | 38.6 | 50.1 | 37.2 | 50.4 | 39.6 | 43.4 | 39.8 | 52.3 | 49.2 | 55.0 | 49.0 | 44.7 | 50.7 | 52.0 |
| bg | 61.3 | – | 38.7 | 39.4 | 39.6 | 34.5 | 46.9 | 25.5 | 26.7 | 42.4 | 22.0 | 43.5 | 29.3 | 29.1 | 25.9 | 44.9 | 35.1 | 45.9 | 36.8 | 34.1 | 34.1 | 39.9 |
| de | 53.6 | 26.3 | – | 35.4 | 43.1 | 32.8 | 47.1 | 26.7 | 29.5 | 39.4 | 27.6 | 42.7 | 27.6 | 30.3 | 19.8 | 50.2 | 30.2 | 44.1 | 30.7 | 29.4 | 31.4 | 41.2 |
| cs | 58.4 | 32.0 | 42.6 | – | 43.6 | 34.6 | 48.9 | 30.7 | 30.5 | 41.6 | 27.4 | 44.3 | 34.5 | 35.8 | 26.3 | 46.5 | 39.2 | 45.7 | 36.5 | 43.6 | 41.3 | 42.9 |
| da | 57.6 | 28.7 | 44.1 | 35.7 | – | 34.3 | 47.5 | 27.8 | 31.6 | 41.3 | 24.2 | 43.8 | 29.7 | 32.9 | 21.1 | 48.5 | 34.3 | 45.4 | 33.9 | 33.0 | 36.2 | 47.2 |
| el | 59.5 | 32.4 | 43.1 | 37.7 | 44.5 | – | 54.0 | 26.5 | 29.0 | 48.3 | 23.7 | 49.6 | 29.0 | 32.6 | 23.8 | 48.9 | 34.2 | 52.5 | 37.2 | 33.1 | 36.3 | 43.3 |
| es | 60.0 | 31.1 | 42.7 | 37.5 | 44.4 | 39.4 | – | 25.4 | 28.5 | 51.3 | 24.0 | 51.7 | 26.8 | 30.5 | 24.6 | 48.8 | 33.9 | 57.3 | 38.1 | 31.7 | 33.9 | 43.7 |
| et | 52.0 | 24.6 | 37.3 | 35.2 | 37.8 | 28.2 | 40.4 | – | 37.7 | 33.4 | 30.9 | 37.0 | 35.0 | 36.9 | 20.5 | 41.3 | 32.0 | 37.8 | 28.0 | 30.6 | 32.9 | 37.3 |
| fi | 49.3 | 23.2 | 36.0 | 32.0 | 37.9 | 27.2 | 39.7 | 34.9 | – | 29.5 | 27.2 | 36.6 | 30.5 | 32.5 | 19.4 | 40.6 | 28.8 | 37.5 | 26.5 | 27.3 | 28.2 | 37.6 |
| fr | 64.0 | 34.5 | 45.1 | 39.5 | 47.4 | 42.8 | 60.9 | 26.7 | 30.0 | – | 25.5 | 56.1 | 28.3 | 31.9 | 25.3 | 51.6 | 35.7 | 61.0 | 43.8 | 33.1 | 35.6 | 45.8 |
| hu | 48.0 | 24.7 | 34.3 | 30.0 | 33.0 | 25.5 | 34.1 | 29.6 | 29.4 | 30.7 | – | 33.5 | 29.6 | 31.9 | 18.1 | 36.1 | 29.8 | 34.2 | 25.7 | 25.6 | 28.2 | 30.5 |
| it | 61.0 | 32.1 | 44.3 | 38.9 | 45.8 | 40.6 | 26.9 | 25.0 | 29.7 | 52.7 | 24.2 | – | 29.4 | 32.6 | 24.6 | 50.5 | 35.2 | 56.5 | 39.3 | 32.5 | 34.7 | 44.3 |
| lt | 51.8 | 27.6 | 33.9 | 37.0 | 36.8 | 26.5 | 21.1 | 34.2 | 32.0 | 34.4 | 28.5 | 36.8 | – | 40.1 | 22.2 | 38.1 | 31.6 | 31.6 | 29.3 | 31.8 | 35.3 | 35.3 |
| lv | 54.0 | 29.1 | 35.0 | 37.8 | 38.5 | 29.7 | 8.0 | 34.2 | 32.4 | 35.6 | 29.3 | 38.9 | 38.4 | – | 23.3 | 41.5 | 34.4 | 39.6 | 31.0 | 33.3 | 37.1 | 38.0 |
| mt | 72.1 | 32.2 | 37.2 | 37.9 | 38.9 | 33.7 | 48.7 | 26.9 | 25.8 | 42.4 | 22.4 | 43.7 | 30.2 | 33.2 | – | 44.0 | 37.1 | 45.9 | 38.9 | 35.8 | 40.0 | 41.6 |
| nl | 56.9 | 29.3 | 46.9 | 37.0 | 45.4 | 35.3 | 49.7 | 27.5 | 29.8 | 43.4 | 25.3 | 44.5 | 28.6 | 31.7 | 22.0 | – | 32.0 | 47.7 | 33.0 | 30.1 | 34.6 | 43.6 |
| pl | 60.8 | 31.5 | 40.2 | 44.2 | 42.1 | 34.2 | 46.2 | 29.2 | 29.0 | 40.0 | 24.5 | 43.2 | 33.2 | 35.6 | 27.9 | 44.8 | – | 44.1 | 38.2 | 38.2 | 39.8 | 42.1 |
| pt | 60.7 | 31.4 | 42.9 | 38.4 | 42.8 | 40.2 | 60.7 | 26.4 | 29.2 | 53.2 | 23.8 | 52.8 | 28.0 | 31.5 | 24.8 | 49.3 | 34.5 | – | 39.4 | 32.1 | 34.4 | 43.9 |
| ro | 60.8 | 33.1 | 38.5 | 37.8 | 40.3 | 35.6 | 50.4 | 24.6 | 26.2 | 46.5 | 25.0 | 44.8 | 28.4 | 29.9 | 28.7 | 43.0 | 35.8 | 48.5 | – | 31.5 | 35.1 | 39.4 |
| sk | 60.8 | 32.6 | 39.4 | 48.1 | 41.0 | 33.3 | 46.2 | 29.8 | 28.4 | 39.4 | 27.4 | 41.8 | 33.8 | 36.7 | 28.5 | 44.4 | 39.0 | 43.3 | 35.3 | – | 42.6 | 41.8 |
| sl | 61.0 | 33.1 | 37.9 | 43.5 | 42.6 | 34.0 | 47.0 | 31.1 | 28.8 | 38.2 | 25.7 | 42.3 | 34.6 | 37.3 | 30.0 | 45.9 | 38.2 | 44.1 | 35.8 | 38.9 | – | 42.7 |
| sv | 58.5 | 26.9 | 41.0 | 35.6 | 46.6 | 33.3 | 46.6 | 27.4 | 30.9 | 38.9 | 22.7 | 42.0 | 28.2 | 31.0 | 23.7 | 45.6 | 32.2 | 44.2 | 32.7 | 31.3 | 33.5 | – |

(using the Acquis corpus) [from Koehn et al., 2009]

Tabella 1: Il-prestazzjonijiet par par miksuba għal 22 lingwa uffiċjali tal-UE fi Traduzzjoni Awtomatika (sors: Euromatrix +)

## It-Teknoloġija Lingwistika 'wara l-kwinti'

Il-bini ta' applikazzjonijiet tat-teknoloġija lingwistika jinvolvi firxa ta' kompiti sekondarji li mhux dejjem jidhru fuq livell ta'

interazzjoni mal-utent, iżda jipprovdu funzjonalitajiet ta' servizz sinifikanti 'wara l-kwinti' tas-sistema. Għaldaqstant, dawn jikkostitwixxu kwistjonijiet importanti ta' riċerka li saru dixxiplini sekondarji individwali tal-Lingwistika Kompjutazzjonali fl-akkademja.

It-tweġib ta' mistoqsijiet sar qasam attiv ta' riċerka, li għalih inbnew il-korpora annotati u bdew kompetizzjonijiet xjentifiċi. L-idea hija li wieħed jimxi minn tfittxija bbażata fuq  kelma ewlenija (li għalih il-magna tirrispondi b'ġabra sħiħa ta' dokumenti potenzjalment rilevanti) għal xenarju fejn l-utent jistaqsi mistoqsija konkreta u s-sistema tipprovdi tweġiba waħda: 'F' liema età Neil Armstrong għamel l-ewwel pass fuq il-qamar?' - '38'. Filwaqt li dan huwa ovvjament relatat ma' Tiftix fuq il-Web tal-qasam ewlieni msemmi qabel, it-tweġib tal-mistoqsijiet, illum huwa primarjament terminu ġenerali għal mistoqsijiet ta' riċerka bħal liema tipi ta' mistoqsijiet għandhom ikunu distinti u kif dawn għandhom ikunu trattati, kif sett ta' dokumenti li potenzjalment fih ir-risposta jista' jiġi analizzat u mqabbel (dawn jagħtu tweġibiet konfliġġenti?), u kif tista' l-informazzjoni speċifika - it-tweġiba – tittieħed b'mod affidabbli minn dokument, mingħajr ma jiġi injorat il-kuntest.

Dan huwa min-naħa l-oħra marbut mal-kompitu tal-estrazzjoni ta' informazzjoni (EI), qasam li kien ferm popolari u influwenti fil-mument tal-'bidla statistika' għal-Lingwistika Kompjutazzjonali, fil-bidu tad-disgħinijiet. EI timmira li tidentifika biċċiet speċifiċi ta' informazzjoni fi klassijiet speċifiċi ta' dokumenti; dan jista' jkun eż. s-sejbien tal-atturi ewlenin fit-teħid ta' kontroll ta' kumpaniji kif irrappurtat fi stejjer fil-gazzetti. Xenarju ieħor li nħadem fuqu huwa rapporti dwar inċidenti terroristiċi, fejn il-problema hija li tqabbel it-test ma' mudell li jispeċifika min hu l-awtur, il-mira, il-ħin u l-post, u r-riżultati tal-inċident. Il-mili tal-mudell permezz ta' dominju speċifiku hija l-karatteristika ċentrali tal-EI, li għal din ir-raġuni huwa eżempju ieħor ta' teknoloġija 'wara l-kwinti' li tikkostitwixxi qasam ta' riċerka demarkata sew, iżda għal skopijiet prattiċi mbagħad jeħtieġ li tkun inkorporata f'ambjent ta' applikazzjoni xierqa.

Żewġ oqsma 'borderline', li xi drabi jkollhom ir-rwol ta' applikazzjoni waħedha u xi kultant dak ta' komponent appoġġjat, 'taħt il-kappa' huma sommarju ta' testi u ġenerazzjoni ta' testi. Sommarju, ovvjament, jirreferi għall-kompitu ta' taqsir ta' test twil, u jiġi offrut pereżempju bħala funzjonalità fi ħdan l-MS Word. Dan jaħdem aktar fuq bażi ta' statistika, billi l-ewwel jidentifika l-kliem 'importanti' f'test (jiġifieri, pereżempju, kliem li huma frekwenti ferm f'dan it-test iżda sostanzjalment anqas frekwenti fl-użu ġenerali tal-lingwa) u mbagħad jiddetermina dawk is-sentenzi li jinkludu ħafna kliem importanti. Dawn is-sentenzi mbagħad jiġu mmarkati fid-dokument, jew estratti minnu, u jiġu meħuda biex isir is-sommarju. F'dan ix-xenarju, li huwa bil-bosta l-aktar wieħed popolari, sommarju huwa ugwali għal estrazzjoni ta' sentenzi: it-test jiġi trattat bħala sett sekondarju tas-sentenzi tiegħu. Is-sistemi ta' sommarji kummerċjali kollha jagħmlu użu minn din l-idea. Metodu alternattiv, li għalih hija ddedikata ċerta riċerka, huwa li sentenzi *ġodda* jiġu attwalment sintetizzati, jiġifieri, jinbena sommarju ta' sentenzi li m'hemmx għalfejn jidhru f'dik il-forma fit-test sors. Dan jeħtieġ ċertu ammont ta' fehim aktar profond tat-test u għaldaqstant huwa ħafna anqas b'saħħtu. Kollox ma' kollox, ġeneratur ta' test huwa f'ħafna każijiet mhux applikazzjoni li tista' toqgħod waħedha iżda huwa inkorporat f'ambjent ta' software

akbar, bħal fis-sistema ta' informazzjoni kliniku fejn dejta tal-pazjent hija miġbura, maħżuna u pproċessata, u l-ġenerazzjoni ta' rapporti hija biss waħda minn ħafna funzjonalitajiet.

## It-Teknoloġija Lingwistika fl-Edukazzjoni

It-teknoloġija lingwistika huwa qasam ferm interdixxiplinarju, li jinvolvi l-kompetenza ta' lingwisti, xjenzjati tal-kompjuter, matematiċi, filosofi, psikolingwisti, u newroxjentisti, fost oħrajn.

F'Malta l-maġġoranza l-kbira ta' riċerka u edukazzjoni fit-TL twettqet fl-Università ta' Malta. Madankollu, din kienet stabbilita pjuttost tard. Raġuni waħda għal dan kienet id-dehra tard tax-Xjenza tal-Kompjuter bħala suġġett kurrikulari fl-Università. It-tmexxija politika turbulenti tal-pajjiż matul l-1970 u l-1980 ma pprevedietx ir-rivoluzzjoni fl-informatika li kellha sseħħ u kien biss fil-bidu tad-disgħinijiet li ġiet offruta għażla ta' kors universitarju permezz tal-Fakultà tax-Xjenza tal-Kompjuter mal-Matematika.

L-għeruq tal-istess bidla seħħew fl-1994, meta twettqet inizjattiva strateġika nazzjonali li tirrikonoxxi u ssaħħaħ ir-rwol tal-TI fis-setturi kummerċjali, politiċi, u fuq kollox, dawk edukattivi. Waħda mill-konsegwenzi immedjati ta' dan kienet l-introduzzjoni ta' programm sostanzjali ta' erba' snin ta' Baċċelerat - il-BSc. IT (Hons) - fl-Università kif ukoll it-twaqqif ta' Dipartiment ġdid tax-Xjenza tal-Kompjuter u Intelliġenza Artifiċjali (CSAI[xxxiii]). Kors fl-NLP kien inkluż bħala għażla avvanzata, u dan wassal, erba' snin wara, għal serje ta' proġetti bħala parti mill-kors għall-ewwel grad fl-aħħar sena tiegħu li trattaw kwistjonijiet tal-ipproċessar tal-lingwa inklużi metodi ta' kompjutazzjoni għall-Malti[xxxiv]. Id-Dipartiment tal-Inġinerija tal-Komunikazzjonijiet u l-Kompjuters ħa sehem ukoll fil-programm, u dan wassal għal sett ieħor ta' proġetti għall-ewwel grad fit-teknoloġija tat-taħdit.

Influwenza oħra importanti fuq ir-riċerka hija L-Istitut tal-Lingwistika tal-Università (IOL), imwaqqaf fl-1988 bil-għan li jgħallem kif ukoll jippromwovi u jikkoordina r-riċerka kemm fil-Lingwistika Ġenerali kif ukoll Applikata, imexxi 'l quddiem ir-riċerka li tinvolvi d-deskrizzjoni ta' lingwi partikolari, mhux l-anqas il-Malti, irawwem l-istudju ta' oqsma sekondarji diversi tal-lingwistika, u jippromwovi riċerka interdixxiplinarja li tinvolvi l-akkademiċi f'kooperazzjoni prattika li tgħaddi bejn konfini dipartimentali u fakultajiet barra mill-pajjiż. L-Istitut tal-Lingwistika jmexxi żewġ programmi għall-ewwel grad: B.A. fil-Lingwistika Ġenerali u l-B.Sc. l-ġdid fit-Teknoloġija Lingwistika Umana li se jkun offrut minn Ottubru 2011. Huwa wkoll possibbli li wieħed jagħmel Masters u Dottorat fil-Lingwistika mal-Istitut.

Fl-1997, grupp interdixxiplinarju ta' xjenzjati tal-kompjuter u lingwisti[xxxv] bdew jaħdmu fuq il-Maltilex, proġett biex jinħoloq dizzjunarju kompjutazzjonali, li kien sostnut minn għotja żgħira mill-Università appoġġjata mill-Mid-Med Bank. Interface sempliċi fuq l-internet kien żviluppat biex jippermetti l-ħolqien u ż-żamma ta' daħliet, kif irrappurtat f'Rosner-et-al (1998) fl-ewwel Grupp ta' Ħidma tal-ACL dwar l-Approċċi Kompjutazzjonali għal Lingwi Semitiċi (Rosner 1998). Eluf ta' daħliet bħal dawn saru b'mod manwali, iżda l-proġett iltaqa' ma' problemi legali, billi l-kumpilazzjoni tad-daħliet kienet fil-biċċa l-kbira mnebbħa mid-dizzjunarju f'forma ta' ktieb ta' Joseph Aquilina (Aquilina 1987).

L-isforz imbagħad ġie trasferit minn dizzjunarji f'forma ta' ktieb għal teħid ta' daħliet lessikali minn sorsi oħra. Żewġ teżijiet (Dalli 2001, Attard 2006) użaw teknika bbażata fuq allinjament meħuda mill-bijoinformatika sabiex jiġbru flimkien daħliet lessikali u dan kien użat bħala mezz ta' strutturar tal-lessiku b'mod awtomatiku.

Minkejja n-nuqqas ta' finanzjament, l-isforz tal-Maltilex issokta b'mod kemxejn frammentat, appoġġjat mill-istaff tal-IOL u d-Dipartiment tas-CSAI. Ma kienx qabel l-2005 li l-Kunsill Malti għax-Xjenza u t-Teknoloġija (MCST) nieda l-ewwel Inizjattiva tar-Riċerka u l-Iżvilupp Teknoloġiku tal-pajjiż u proposta konġunta għas-Server għar-Riżorsi Lingwistiċi bil-Malti (MLRS) kienet aċċettata, sakemm ikun hemm appoġġ finanzjarju suffiċjenti biex jimpjegaw riċerkatur full time bejn l-2006 u l-2008. Il-proġett kellu żewġ miri kemm li joħloq dizzjunarju kif ukoll korpus (Rosner 2008), u stabbilixxa l-pedamenti għas-server tal-MLRS preżenti.

Ir-riċerka msemmija hawn fuq tittratta prinċipalment mal-lingwa miktuba. Żewġ fergħat tax-xogħol relatat mat-taħdit ukoll qegħdin jitwettqu.

L-ewwel waħda, mibdija minn tradizzjoni ta' pproċessar tas-sinjali fi ħdan il-Fakultà tal-Inġinerija, ħolqot prototip ta' sintetizzatur għat-taħdit (Micallef 1997). Ix-xogħol tiegħu influwenza diversi proġetti oħra mmirati biex itejbu s-sinteżi tat-taħdit minn perspettiva baxxa ta' riżorsi inklużi Calleja (2002), Farrugia (2005), Camilleri (2010), Borg-et-al (2011).

It-tieni, tittratta l-kwistjoni tal-intonazzjoni (Vella 2007) minn perspettiva lingwistika. Xi xogħol pijunier biex jinħoloq korpus u qafas deskrittiv għall-istudju tal-intonazzjoni tal-Malti sar minn Vella u Farrugia (Vella u Farrugia 2006).

Barra minn Malta, żewġ gruppi ta' riċerka li qegħdin f'kollaborazzjoni attiva ma' sforzi lokali mmirati lejn it-TL jistħoqqilhom aċċenn speċjali.

Fl-Università ta' Arizona, grupp immexxi mil-lingwista Adam Ussishkin huwa partikolarment interessat fil-kwistjonijiet psikolingwistiċi li jappartjenu għal-lingwi semitiċi inkluż il-Malti. Sabiex jiġu studjati dawn il-kwistjonijiet sar disponibbli korpus online (Ussishkin et-al 2009).

Fl-Università ta' Bremen, il-Professur Thomas Stolz kien involut b'mod attiv fl-istudju akkademiku tal-Malti iżda huwa partikolarment magħruf talli ospita l-ewwel konferenza dwar il-Lingwistika Maltija fi Bremen (Comrie et al 2009), waqqaf ġurnal[xxxvi] u l-Għaqda Internazzjonali tal-Lingwistika Maltija, ibbażata wkoll fi Bremen, li taħdem flimkien mal-Kunsill għall-Ilsien Malti ibbażat f'Malta.

Kif diġà ssemma, il-komunitajiet sensittivi għat-TL li hemm fl-Università ta' Malta jinsabu prinċipalment fi ħdan il-Fakultà tal-ICT, l-Istitut tal-Lingwistika. Hemm ukoll interess potenzjali fil-Fakultà tal-Arti (Dipartiment tal-Malti) u suġġetti oħra Umanistiċi għalkemm sa issa hemm it-tendenza li l-lingwistika kompjutazzjonali titqies bħala suġġett eżotiku li jinsab f'fakultajiet aktar xjentifiċi bħax-xjenza tal-kompjuter jew l-istudji umanistiċi u, għalhekk, it-temi ta' riċerka li ġew trattati jikkoinċidu b'mod parzjali biss.

Ħaġa kurjuża, Malta mhijiex nieqsa minn avvenimenti internazzjonali relatati mat-TL. L-LREC 2010 saret fil-Belt Valletta, u ġibdet 1200 parteċipant. Il-konferenza annwali tal-EAMT seħħet ukoll f'Malta fl-1994, u saru wkoll numru ta' workships iżgħar matul dawn l-aħħar 10 snin.

## Programmi ta' Teknoloġija Lingwistika

Malta ssieħbet fl-UE fl-2004 u dan l-avveniment immedjatament ta lill-Malti l-istatus ta' lingwa uffiċjali tal-UE. Flimkien ma' dan l-istatus inħolqu obbligi ġodda – b'mod partikolari biex jiġu tradotti kwantitajiet kbar ta' dokumenti uffiċjali, u barra minn hekk, ir-rikonoxximent, fuq livell Ewropew, li bħala lingwa nazzjonali, għandu jkollu status tal-"ewwel klassi" minn perspettiva teknoloġika kif ukoll soċjali, u jingħata d-drittijiet u privileġġi kollha li jgawdu l-"akbar" lingwi Ewropej (jiġifieri li għandhom numri akbar ta' kelliema nattivi).

L-Istrateġija Nazzjonali tat-TI 2008-10 tal-gvern inkludiet numru ta' għanijiet marbutin mal-Lingwa Maltija inkluż (i) l-iżvilupp tal-gvern fuq l-internet bil-Malti, (ii) il-ħolqien ta' għodod għal-lingwa Maltija, b'kollaborazzjoni mal-Università, u (iii) appoġġ għal komunitajiet fuq l-internet bil-Malti. Waqt li qed jinkiteb dan fl-2011, mhux l-għanijiet kollha ntlaħqu. Madankollu l-effetti fit-tul ta' din l-istrateġija qed jibdew jieħdu forma.

Bħalissa x-xena tat-teknoloġija lingwistika f'Malta tinsab taħt l-influwenza ta' erba' inizjattivi prinċipali:

1 L-ewwel nett, proġett appoġġjat mill-gvern parzjalment iffinanzjat mill-fond ta' żvilupp reġjonali tal-UE qiegħed fil-proċess li jwassal li t-teknoloġija tat-taħdit tkun tista' tintlaħaq minn persuni b'diżabiltà. Il-proġett bħalissa qed jiffoka fuq sinteżi tat-taħdit bil-Malti, u f'dan il-punt il-mudelli tal-lingwa rilevanti qegħdin fil-proċess li jiġu żviluppati. Il-konsorzju, li jikkonsisti f'SME[xxxvii], fondazzjoni[xxxviii], u l-Università, wiegħed li dawn ir-riżorsi se jkunu disponibbli għal skopijiet ta' riċerka. Wieħed għad irid jara jekk il-komponenti tas-sintetizzatur tat-taħdit se jkun disponibbli għan-networks li jqassmu r-riżorsi ispirati minn CLARIN u META.It-tieni nett, kif jirriżulta mir-rapport kurrenti, Malta qed tipparteċipa fil-METANET4U u għalhekk tirċievi finanzjament sinifikanti mill-UE mmirat lejn it-tisħiħ u d-distribuzzjoni ta' riżorsi u għodod li huma speċifikament għall-Malti. L-Università ta' Malta hija membru tal-META-NET u l-ħsieb huwa li twettaq l-obbligi tagħha lejn l-għanijiet tal-META, partikolarment rigward l-identifikazzjoni tal-partijiet interessati, attwali u potenzjali.

3 It-tielet, is-Server għar-Riżorsi Lingwistiċi bil-Malti (MLRS) qed jagħti l-frott u sforzi sinifikanti għadhom għaddejjin fl-Università, permezz tal-Istitut tal-Lingwistika[xxxix] u d-Dipartiment ta' Sistemi Intelliġenti tal-Kompjuter[xl], li jsostnu u jiżviluppaw dan. Bħalissa l-MLRS huwa online fuq http://mlrs.research.um.edu.mt. Il-korpus jinkludi xi 80m kelma, u s-sistema tinkludi xi servizzi bażiċi li jinkludu tiftix KWIC u stampi, tfittxija skont il-mudelli, diversi tipi ta' analiżi statistika eċċ. Bħalissa hemm aktar għodod ippjanati inkluż tagger għall-kategoriji tal-kliem u ċekkjatur ortografiku.

4. Fl-aħħarnett, programm ġdid għall-ewwel grad fit-Teknoloġija tal-Lingwa Umana għandu jiġi mniedi mill-Istitut tal-Lingwistika f'Ottubru 2011. Dan se jkopri firxa sħiħa ta' suġġetti u inevitabilment se jħalli impatt pożittiv fit-tul fuq l-istudju tal-Malti minn perspettiva kompjutazzjonali.

Minbarra dawn, proġett biex tiġi żviluppata verżjoni elettronika tad-dizzjunarju ta' Aquilina (Aquilina 1987) qed titħejja bħalissa. Dan huwa sforz kollaborattiv bejn l-Università ta' Malta li qed tforni l-kompetenza lingwistika, l-Università ta' Arizona, li diġà ddiġitalizzaw id-dizzjunarju f'forma li tinqara minn kompjuter, u l-pubblikaturi Midsea Books Valletta. L-għanijiet doppji tal-proġett huma li jiġi aġġornat il-kontenut, u biex jagħtu lir-riċerkaturi l-flessibilità biex jaċċessaw it-test malajr. Għaddej sforz lokalment, sabiex jiġi organizzat livell tajjeb ta' kompetenza lessikografika meħtieġa għall-aġġornament tal-kontenut.

Għandna wkoll insemmu r-relazzjoni ta' Malta mal-CLARIN, proposta ta' infrastruttura ta' riċerka tal-UE li tindirizza l-provvista ta' riżorsi lingwistiċi għax-Xjenzi Umanistiċi u Soċjali. Matul il-fażi ta' speċifikazzjoni, l-Università setgħet tipparteċipa bis-saħħa ta' għotja żgħira ta' appoġġ mill-Kunsill lokali għax-Xjenza u t-Teknoloġija. Madankollu, l-isfida biex jiġi żgurat il-finanzjament fit-tul meħtieġ għall-fażi ta' kostruzzjoni tal-CLARIN kienet akbar. L-identifikazzjoni ta' entità tal-gvern xierqa biex tieħu r-responsabbiltà għall-programm s'issa kienet mingħajr suċċess. Konsegwentement, il-parteċipazzjoni futura ta' Malta fil-fażi ta' kostruzzjoni s'issa għadha mhix deċiża.

## Disponibbiltà ta' Għodod u Riżorsi għall-Malti

It-tabella li ġejja tipprovdi ħarsa ġenerali lejn is-sitwazzjoni kurrenti tal-appoġġ tat-teknoloġija lingwistika għall-Malti. Il-klassifikazzjoni ta' teknoloġiji eżistenti u riżorsi hija bbażata fuq stimi studjati minn esperti ewlenin diversi li użaw dawn il-kriterji li ġejjin (kull waħda tvarja minn 0 sa 6).

1. **Kwantità:** Teżisti għodda / riżors għal-lingwa msemmija? Aktar ma hemm għodod / riżorsi eżistenti, aktar hija għolja l-klassifikazzjoni.

   - 0: ma hemm assolutament l-ebda għodod / riżorsi
   - 6: ħafna għodod / riżorsi, varjetà kbira

2. **Disponibiltà:** L-għodod / ir-riżorsi huma aċċessibbli, jiġifieri, huma Sors Miftuħ, jintużaw b'mod liberu fuq kull pjattaforma jew disponibbli biss għal prezz għoli jew taħt kundizzjonijiet ristretti ħafna?

   - 0: prattikament l-għodod / ir-riżorsi kollha huma disponibbli għal prezz għoli biss
   - 6: ammont kbir ta' għodod / riżorsi huma disponibbli liberament u bil-miftuħ taħt Sors Miftuħ sensibbli jew Liċenzji tal-Creative Commons li jippermettu kemm l-użu kif ukoll tibdil ta' format mill-ġdid

3. **Kwalità:** Kemm qed jintlaħqu tajjeb il-kriterji rispettivi tal-prestazzjoni ta' għodod u indikaturi ta' kwalità tar-riżorsi permezz tal-aqwa għodod, applikazzjonijiet jew riżorsi

disponibbli? Dawn l-għodod / riżorsi kurrenti qed issirilhom ukoll manutenzjoni b'mod attiv?

□0: riżors / għodda ta' kwalità bażika

□6: għodda ta' kwalità għolja, annotazzjonijiet ta' kwalità umana f'riżorsa

4 **Kopertura:** F'liema grad l-aqwa għodod jilħqu l-kriterji ta' kopertura rispettivi (stili, ġeneri, xorta ta' testi, fenomeni lingwistiċi, tipi ta' input/output, numru ta' lingwi appoġġjati minn sistema TA eċċ)? F'liema grad qegħdin ir-riżorsi rappreżentattivi tal-lingwa fil-mira jew lingwi sekondarji?

□0: riżors b'għan speċjali jew għodda, każ speċifiku, kopertura żgħira ħafna, jistgħu jintużaw biss għal każijiet speċifiċi ħafna, ta' użu mhux ġenerali

□6: riżors b'kopertura wiesgħa ħafna, għodda qawwija ħafna, ġeneralment applikabbli, appoġġjat b'ħafna lingwi

5 **Maturità:** Jistgħu l-għodda/ r-riżors jitqiesu matura, stabbli, lesti għas-suq? L-aqwa għodod/riżorsi disponibbli jistgħu jintużaw kif inhuma jew iridu jiġu adattati? Il-prestazzjoni ta' din it-teknoloġija hija xierqa u lesta għall-użu ta' produzzjoni jew hija biss prototip li ma jistax jintuża għal sistemi ta' produzzjoni? Indikatur jista' jkun jekk ir-riżorsi/l-għodod humiex aċċettati mill-komunità u użati b'suċċess fis-sistemi TL.

□0: prototip preliminari, sistema ta' kwalità bażika, prova tal-kunċett, eżempju ta' eżerċizzju tar-riżorsi

□6: komponent li jista' jiġi integrat/applikat immedjatament

6 **Sostenibbiltà:** L-għodda/riżors kemm jistgħu jinżammu /jiġu integrati b'mod tajjeb fis-sistemi preżenti tat-TI? L-għodda/ir-riżors jissodisfaw ċertu livell ta' sostenibbiltà fir-rigward ta' dokumentazzjoni/manwali, spjegazzjoni tal-każijiet ta' użu, il-bidu tas-sistema, GUIs eċċ? Hija tuża prattika standard/l-aħjar prattika ta' ambjenti ta' programmar (bħal Java EE)? Industrija/standards ta' riċerka/kważi standards jeżistu u jekk iva, l-għodda/r-riżors huma konformi (formati ta' dejta eċċ)?

□0: formati ta' dejta *ad hoc* u APIs kompletament proprjetarji

□6: konformità sħiħa mal-istandard, dokumentazzjoni sħiħa

7 **Adattabiltà:** L-aqwa għodod jew riżorsi kemm jistgħu jiġu adattati/estiżi tajjeb għal każijiet ta' kompiti/ dominji/ġeneri/tipi ta' testi/użi ġodda eċċ?

□0: prattikament impossibbli li wieħed jadatta għodda/riżors għal kompitu ieħor, impossibbli anki ma' ammonti kbar ta' riżorsi jew xhur ta' persuni disponibbli

□6: livell għoli ħafna ta' adattabiltà; adattament faċli ħafna wkoll u possibbli b'mod effiċjenti

## Tabella ta' Għodod u Riżorsi għall-Malti

| | Kwantità | Disponibiltà | Kwalità | Kopertura | Maturità | Sostenibiltà | Adattabiltà |
|---|---|---|---|---|---|---|---|
| **Teknoloġija Lingwistika (Għodod, Teknoloġiji, Applikazzjonijiet)** | | | | | | | |
| **Sistema ta' tokens, Morfoloġija** (sistema ta' tokens, POS tagging, analiżi/ġenerazzjoni morfoloġika) | 2 | 4 | 3 | 4 | 2 | 3 | 3 |
| **Parsing** (analiżi sintattika bażika jew fid-dettal) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Semantika ta' Sentenzi** (WSD, struttura ta' argumenti, rwoli semantiċi) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Semantika ta' Testi** (riżoluzzjoni tal-koreferenza, kuntest, prammatika, inferenza) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Ipproċessar ta' Diskors Avvanzat** (struttura ta' testi koerenza, struttura retorika/RST, taqsim f'zoni argumentattivi, argumentazzjoni, mudelli ta' testi, tipi ta' testi eċċ.) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Rkupru ta' Informazzjoni** (indiċjar ta' testi, RI multimidjali, RI bejn il-lingwi) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Estrazzjoni ta' Informazzjoni** (identifikazzjoni ta' entità msemmija, estrazzjoni tal-okkorrenza/tar-relazzjoni, identifikazzjoni tal-opinjoni/ tas-sentiment, minjiera/analitika ta' testi) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Ġenerazzjoni ta' Lingwa** (ġenerazzjoni ta' sentenzi, ġenerazzjoni ta' rapporti, ġenerazzjoni ta' testi) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Sommarju, Tweġib ta' Mistoqsijiet, Teknoloġiji Avvanzati ta' Aċċess għal Informazzjoni** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Traduzzjoni Awtomatika** | 2 | 2 | 2 | 3 | 2 | 2 | 2 |
| **Identifikazzjoni ta' Taħdit** | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Sinteżi ta' Taħdit** | 3 | 1 | 4 | 4 | 3 | 3 | 3 |
| **Ġestjoni ta' Djalogu** (kapaċitajiet ta' djalogu u mmudellar ta' utent) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Riżorsi Lingwistiċi (Riżorsi, Dejta, Bażijiet ta' Għarfien)** | | | | | | | |
| **Korpora ta' Referenza** | 4 | 4 | 3 | 3 | 3 | 4 | 4 |

| | Kwantità | Disponibiltà | Kwalità | Kopertura | Maturità | Sostenibiltà | Adattabiltà |
|---|---|---|---|---|---|---|---|
| **Korpora ta' Sintassi** (banek ta' siġar, banek ta' dipendenza) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Korpora ta' Semantika** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Korpora ta' Diskors** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Korpora Paralleli, Memorji tat-Traduzzjoni** | 4 | 4 | 3 | 2 | 2 | 2 | 2 |
| **Korpora ta' Taħdit** (dejta ta' taħdit mhux ipproċessata, dejta ta' taħdit immarkata/annotata, dejta ta' djalogu ta' taħdit) | 3 | 1 | 3 | 2 | 3 | 3 | 2 |
| **Dejta multimidjali u multimodali** (dejta ta' testi flimkien ma' audio/video) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Mudelli ta' Lingwi** | 2 | 1 | 3 | 3 | 3 | 1 | 1 |
| **Dizzjunarji, Terminoloġiji** | 3 | 3 | 2 | 3 | 3 | 3 | 3 |
| **Grammatiki** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Teżawri, WordNets** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Riżorsi Ontoloġiċi għal Għarfien Dinji** (mudelli fi grad għoli eż. Dejta Llinkjata) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

## Konklużjonijiet

F'din is-Serje ta' White Papers, sar l-ewwel sforz biex tiġi evalwata s-sitwazzjoni globali ta' ħafna lingwi Ewropej f'dak li jirrigwarda l-appoġġ tat-teknoloġija lingwistika b'mod li jippermetti tqabbil ta' livell għoli u identifikazzjoni ta' nuqqasijiet u ħtiġijiet.

Għall-Malti, il-karatteristiċi l-aktar evidenti li ħarġu mit-tabella huma li

- l-biċċa l-kbira tad-daħliet huma vojta, u
- l-ogħla grad li ntlaħaq huwa 4.

Il-fatt li d-daħliet huma kważi kollha vojta jirrifletti l-istat immatur ta' riċerka u żvilupp marbut mat-TL f'Malta. Għalkemm hemm sinjali li s-sitwazzjoni qed titjieb, l-investiment fit-teknoloġija lingwistika jibqa' fuq livell baxx, u bħala riżultat, minkejja l-kisbiet lokali modesti, l-isforz huwa frammentat, kemm f'termini ta' kopertura ta' oqsma differenti, kif ukoll f'termini ta' sostenibbiltà

ta' riċerka: kien hemm wisq proġetti li jinvolvu qasam wieħed, riċerkatur wieħed biss, u sena jew sentejn biss. L-isforzi kollettivi ma jammontawx għal dak li huwa mixtieq.

Allura x'inkiseb? Nistgħu naraw billi nħarsu lejn id-daħliet mhux vojta, li l-medja tal-punteġġ tagħhom jagħti l-ordni li ġejja:

- Għodod:
    1. Sistema ta' tokens, Sinteżi ta' Taħdit
    2. Identifikazzjoni ta' Taħdit
- Riżorsi:
    1. Korpora ta' Referenzi
    2. Korpora Paralleli
    3. Dizzjunarji, Terminoloġiji (huwa mifhum li dawn jinkludu listi ta' kliem)
    4. Mudelli lingwistiċi

Fir-rigward tal-għodod:

Estrazzjoni ta' testi f'livell baxx u għodod ta' pproċessar huma disponibbli, inkluż fornitur ta' tokens. POS-tagger qed jiġi żviluppat, iżda l-prestazzjoni tiegħu mhijiex l-aktar stat modern, sakemm ikun hemm aktar taħriġ b'dejta annotata aħjar.

Għodod ta' livell ogħla (analiżi sintattika jew semantika, għodod ta' klassifikazzjoni, estrazzjoni ta' informazzjoni eċċ.) huma kompletament neqsin. Il-konsegwenza hija li, pereżempju, m'hemm l-ebda treebanks disponibbli għall-Malti.

Prototip ta' għodod ta' identifikazzjoni ta' taħdit ġew żviluppati fl-Università iżda mhumiex faċilment disponibbli fiż-żmien meta qed jinkiteb dan. Madankollu, il-magna tat-taħdit iffinanzjata mill-gvern imsemmija qabel għandha tipprovdi sintetizzatur tat-taħdit jiffunzjona sal-2013. Filwaqt li dan huwa żvilupp pożittiv ħafna, huwa ffukat ħafna fuq in-naħa tas-sinteżi tat-taħdit. Kważi l-ebda xogħol fuq l-identifikazzjoni tat-taħdit ma huwa ppjanat f'dan l-istadju.

Fir-rigward ta' riżorsi, is-sitwazzjoni hija xi ftit aktar strutturata, minħabba li diġà jeżisti l-MLRS, infrastruttura kompjutazzjoni estensiva fil-forma ta' server li tipprovdi funzjonalità bażika li tippermetti aċċess fuq il-web għall-korpora disponibbli, xi servizzi, u sistema rudimentali li tiffaċilita l-preżentazzjoni ta' kontribuzzjonijiet. L-MLRS bħalissa qed jipprovdi xi servizzi bażiċi ħafna għall-estrazzjoni, rappreżentazzjoni, tiftix u analiżi ta' testi.

Il-korpus tal-MLRS eżistenti bħalissa għandu kobor ta' madwar 100 miljun tokens. Dan huwa prinċipalment testwali u monolingwali. Huwa wkoll kemxejn mhux rappreżentattiv: hemm abbundanza ta' materjal legalistiku, iżda nuqqas ta' testi akkademiċi u xogħlijiet fittizji.

F'dan l-istadju, dan il-materjal jista' biss jiġi mfittex u analizzat permezz tas-server u ma jistax jiġi aċċessat direttament. Ir-raġunijiet huma legalistiċi. B'aċċess ristrett b'dan il-mod, il-kumplikazzjonijiet tal-IPR u d-drittijiet tal-awtur ġew evitati bil-pulit. Il-prezz huwa li dawn il-kumplikazzjonijiet eventwalment se jkollhom jiġu kkonfrontati fil-futur, u fil-fatt META tinsab fil-proċess ta' formulazzjoni ta' sett ta' ftehim ta' liċenzjar li jgħodd għad-distribuzzjoni tar-riżorsi, bħall-MLRS.

F'dan ir-rapport, ippruvajna nwasslu l-istat paradossali tat-Teknoloġija Lingwistika tal-Malti. Il-paradoss jinħoloq għaliex hemm sforzi sinifikanti minn numru żgħir ta' nies ikkwalifikati sew f'firxa ta' attivitajiet marbuta mat-TL biex itejbu l-istat tat-teknoloġija, jew f'termini ta' għodod, jew ta' riżorsi, jew it-tnejn li huma. Huwa wkoll ċar li fil-kuntest usa' ta' attivitajiet edukattivi, kummerċjali u kulturali fil-pajjiż, hemm post għat-TL biex tagħti kontribut importanti. Il-problema hija li l-isforzi li saru ma kinux koordinati, kienu fuq żmien qasir, u frammentarji, u għaldaqstant il-progress huwa aktar bil-mod milli seta' jkun.

Koordinazzjoni sostnuta u diretta tal-isforz hija, fl-opinjoni tagħna, l-uniku mod kif il-benefiċċji tat-TL għall-Malti jistgħu jiġu rrealizzati fi żmien raġonevoli. Aħna nemmnu li anki f'pajjiż żgħir bħal Malta, ix-xogħol jeħtieġ li jitqassam fost partijiet interessati differenti. Irridu naslu għal pjan prattikabbli permezz ta' verżjoni lokalizzata tad-diviżjoni tax-xogħol bejn tliet partijiet, li hija sostnuta mill-META: identifikazzjoni ta' komunità b'viżjoni komuni; estensjoni ta' infrastruttura li tiffaċilita t-tqassim ta' riżorsi, u tisħiħ tal-konnessjonijiet bejn it-TL u l-oqsma ġirien ta' riċerka u żvilupp.

## Referenzi


Acquilina Guze, Maltese-English Dictionary, Valletta: Midsea Publications, 1987.

Attard, D., A Lexicon Server Toolkit for Maltese, Dept CSAI, L-Università ta' Malta, 2005

Bajada, J., Investigation of Translations Equivalences from Parallel Texts, Dept CSAI, L-Università ta' Malta, 2004

Bajada, J., Phrase Extraction for Machine Translation, MSc, Dept CSAI, L-Università ta' Malta, 2009.

Buhagiar, I, and Micallef, P., Web Based Maltese Language Text to Speech Synthesiser, Proceedings of WICT08, L-Università ta' Malta, 2008

Bonnici, S., A Hypertext Implementation of a Dictionary for the Maltese Language, Dept. CSAI, L-Università ta' Malta, 1997

Borg, M., K Bugeja, C Vella, G Mangion, C Gafa, Preparation of a free-running text corpus for Maltese concatenative speech synthesis, GĦILM 3rd Conference on Maltese Linguistics, Valletta, 2011.

Calleja, S., Speech Synthesis, MSc, Dept CSAI, L-Università ta' Malta, 2002.

Camilleri, R., Speech Annotation System, Roberta Camilleri, B.Eng Project, Dept Computer Communications Engineering, Teżi tal-M.Sc, 2010.

Comrie, B., Fabri, R., Mifsud, M., Stolz, T., & Vanhove, M. (Eds.). Introducing Maltese Linguistics. Proceedings of the 1st International conference on Maltese Linguistics (Bremen/Germany, October, 2007). Studies in Language Companion Series. Amsterdam ; Philadelphia: John Benjamins, 2009

Dalli, A., Computational Lexicon for Maltese, MSc, Dept CSAI, L-Università ta' Malta, 2001.

Galea, David, A System for the Analysis of Maltese Verbs, Dept CSAI, L-Università ta' Malta, 1999.

Farrugia, P., An Automatic Translation System for Maltese/English, Dept CSAI, L-Università ta' Malta, 1999.

Farrugia, P., Text to Speech Technologies for Mobile Telephony Services, CSAI, L-Università ta' Malta, 2005.

Farrugia, A., MULTIMORPH: A Computational Analysis of the Maltese Broken Plural, Dept CSAI, L-Università ta' Malta, 2008

Farrugia, C., A Portal for Acquisition, Representation and Presentation of Current Affairs in Malta, Dept CSAI, L-Università ta' Malta, 2009

Farrugia, R, SAMILS – A Semi-Automatic Machine Indexing for Legal Systems, Dept CSAI, L-Università ta' Malta, 2000

Felter, P. an Optical Character Recognition System for Maltese, Dept CSAI, L-Università ta' Malta, 2001

Mangion, Gordon, Spelling Correction for Maltese, Dept CSAI, L-Università ta' Malta, 1999.

Micallef, P., A Text to Speech Synthesis System for Maltese, Teżi Dottorat, L-Università ta' Surrey, Diċembru 1997.

Mizzi, R., The Development of a Statistical Spell Checker for Maltese, Dept CSAI, L-Università ta' Malta, 2000.

Psaila, A. Speech Annotation using Hidden Markov Models, Dept Computer Communications Engineering, Teżi tal-M.Sc, 2008

Rosner, M., J. Caruana, u R. Fabri. Maltilex: A computational lexicon for Maltese. F'M. Rosner, editur, Computational Approaches to Semitic Languages: Proceedings of the Workshop held at COLING-ACL98, Université de Montréal, Canada, pp 97–105, 1998.

Rosner, M., R. Fabri, D. Attard, u Albert Gatt. Maltese language resource server. In Proceedings of CSAW06, L-Università ta' Malta, paġni 90–98, November 2006.

Rosner. M., Electronic language resources for Maltese. Fi Proceedings of Bremen Workshop on Maltese Linguistics, L-Università ta' Bremen, 2008.

Sciriha, L., The Maltese Interactive Picture Dictionary, Protea Textware, ISBN 978-0958733007, 1997.

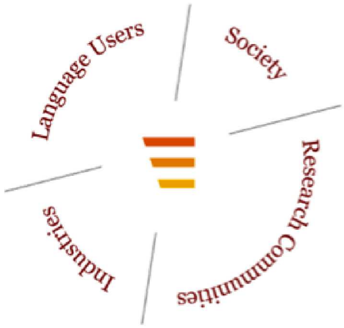
Ussishkin, A., Francom, J.,Woudstra, D., Creating a Web-based Lexical Corpus and Information-extraction Tools for the Semitic Language Maltese, Proc SALTMIL Workshop, Donostia, ISBN: 978-84-692-4940-6Settembru 2009.

Vella A. u P. Farrugia, MalToBI – Building an Annotated Corpus of Spoken Maltese, 2006.

Vella, A., On Maltese Prosody or The Intonational Phonology of Maltese, Proc. 2nd International Conference of, Maltese Linguistics, Bremen, Ottubru 2007

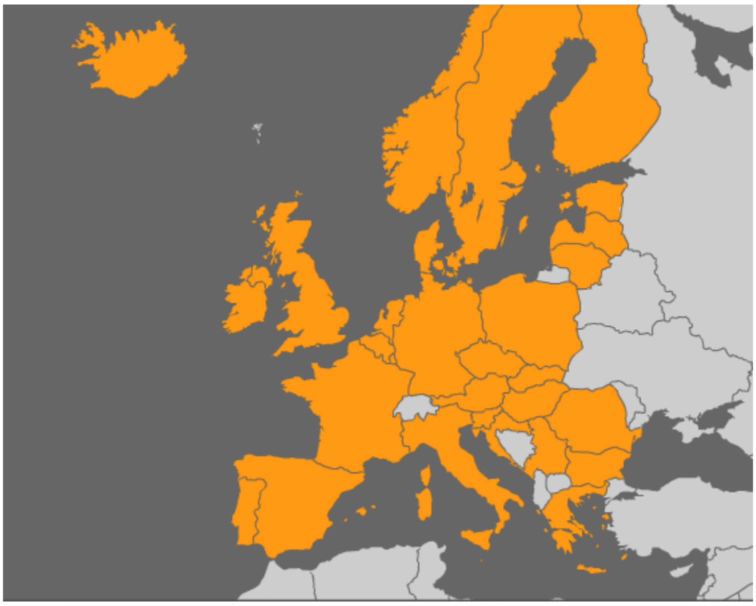
Vella, G., Automatic Summarization of Legal Documents, Dept. CSAI, L-Università ta' Malta, 2010

# Dwar META-NET




*The Multilingual Europe Technology Alliance (META)*

META-NET huwa Network ta’ Eċċellenza ffinanzjat mill-Kummissjoni Ewropea. In-network bħalissa jikkonsisti f’47 membru minn 31 pajjiż Ewropew. META-NET irawwem Alleanza Ewropea ta’ Teknoloġija Multilingwi (META), komunità dejjem tikber ta’ professjonisti u organizzazzjonijiet tat-teknoloġija lingwistika fl-Ewropa.



Grafika 1: Pajjiżi Rappreżentati fil-META-NET

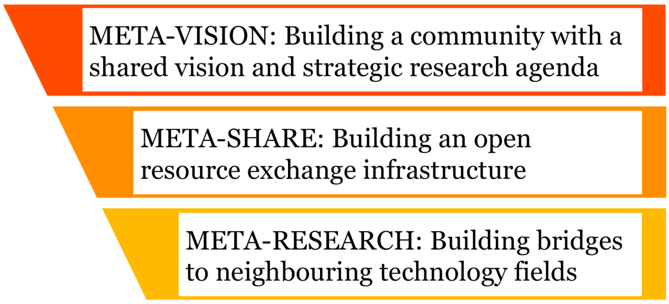
META-NET jikkopera ma’ inizjattivi oħra bħal Riżorsi Lingwistiċi Komuni u Infrastruttura Teknoloġika (CLARIN), li qed jgħinu jistabbilixxu riċerka umanistika diġitali fl-Ewropa. META-NET irawwem is-sisien teknoloġiċi għall-istabbiliment u ż-żamma ta’ soċjetà tal-informazzjoni Ewropea tassew multilingwi:

- jagħmel il-komunikazzjoni u l-kooperazzjoni bejn il-lingwi possibbli;
- jipprovdi aċċess ugwali għall-informazzjoni u l-għarfien fi kwalunkwe lingwa;
- joffri teknoloġija tal-informatika f’network avvanzat u affordabbli għaċ-ċittadini Ewropej.

META-NET jistimula u jippromwovi teknoloġiji multilingwi għal-lingwi Ewropej kollha. It-teknoloġiji jippermettu traduzzjoni awtomatika, produzzjoni tal-kontenut, ipproċessar ta’ informazzjoni u ġestjoni ta’ għarfien għal varjetà wiesgħa ta’ applikazzjonijiet u dominji ta’ suġġetti. In-network irid li jitjiebu l-metodi kurrenti, sabiex tkun tista’ sseħħ komunikazzjoni u kooperazzjoni aħjar bejn l-ilsna. L-Ewropej għandhom dritt ugwali għall-informazzjoni u l-għarfien irrispettivament mil-lingwa.

## Linji ta' azzjoni

META-NET tnieda fl-1 ta' Frar 2010 bil-għan li javvanza r-riċerka fit-teknoloġija lingwstika (LT). In-network jappoġġja Ewropa li tingħaqad bħala suq diġitali u spazju ta' informazzjoni uniku. META-NET wettaq diversi attivitajiet li jippromwovu l-għanijiet tiegħu. META-VISION, META-SHARE u META-RESEARCH huma t-tliet linji ta' azzjoni tan-network.




Grafika 2: Tliet Linji ta' Azzjoni fil-META-NET

**META-VISION** trawwem komunità dinamika u influwenti ta' partijiet interessati, li tingħaqad madwar viżjoni komuni u aġenda ta' riċerka strateġika komuni (SRA). L-enfasi ewlenija ta' din l-attività hija li tinbena komunità tat-TL koerenti u koeżiva fl-Ewropa billi jinġabru flimkien rappreżentanti minn gruppi frammentati ħafna u diversi ta' partijiet interessati. Fl-ewwel sena tal-META-NET, preżentazzjonijiet fil-Forum ta' FLaReNet (Spanja), Language Technology Days (Lussemburgu), JIAMCATT 2010 (Lussemburgu), LREC 2010 (Malta), EAMT 2010 (Franza) u l-ICT 2010 (il-Belġju) iċċentraw fuq l-involviment tal-pubbliku. Skont stimi inizjali, META-NET diġà kkuntattja aktar minn 2,500 professjonist tat-TL biex magħhom jiżviluppa l-għanijiet u l-viżjonijiet tiegħu. Fl-avveniment META-FORUM fi Brussell tal-2010, META-NET ikkomunika r-riżultati inizjali tal-viżjoni tiegħu li jibni proċess għal aktar minn 250 parteċipant. F'serje ta' sessjonijiet interattivi, il-parteċipanti pprovdew feedback fuq il-viżjonijiet ppreżentati min-network.

**META-SHARE** toħloq faċilità miftuħa u mqassma għal skambju u tqassim ta' riżorsi. In-network 'peer to peer' ta' repożitorji se jinvolvi dejta lingwistika, għodod u servizzi tal-web li huma dokumentati bi kwalità għolja ta' metadejta u organizzati f'kategoriji standardizzati. Ir-riżorsi jridu jkunu faċilment aċċessibbli u mfittxija b'mod uniformi. Ir-riżorsi disponibbli jinkludu materjal miftuħ u ħieles ta' sorsi kif ukoll oġġetti ristretti, kummerċjalment disponibbli, ibbażati fuq tariffi. META-SHARE timmira lejn dejta eżistenti lingwistika, għodod u sistemi kif ukoll prodotti ġodda li qed jitfaċċaw u li huma meħtieġa għall-bini u l-evalwazzjoni ta' teknoloġiji, prodotti u servizzi ġodda. L-użu mill-ġdid, it-tagħqid flimkien, l-użu f'format differenti u inġinerija mill-ġdid ta' dejta lingwistika u għodod għandu rwol kruċjali. META-SHARE eventwalment se ssir parti importanti mis-suq tat-TL għall-iżviluppaturi, l-esperti ta' lokalizzazzjoni, riċerkaturi, tradutturi u professjonisti tal-lingwa minn intrapriżi żgħar, ta' daqs medju u kbar. META-SHARE tindirizza ċ-ċiklu ta' żvilupp sħiħ tat-TL – minn riċerka għal prodotti u servizzi innovattivi. Aspett ewlieni ta' din l-attività huwa li jistabbilixxi lil META-SHARE bħala

parti importanti u prezzjuża ta' infrastruttura Ewropea u globali għall-komunità tat-TL.

**META-RESEARCH** tibni pontijiet għal oqsma marbuta mat-teknoloġija. Din l-attività tfittex li tmexxi l-avvanzi f'oqsma oħra u tikkapitalizza fuq riċerka innovattiva li tista' tibbenefika t-teknoloġija lingwistika. B'mod partikolari, din l-attività trid tinkludi aktar semantika fit-traduzzjoni awtomatika (TA), twettaq l-aħjar tqassim ta' xogħol ibridu tat-TA, tisfrutta l-kuntest waqt kompjutazzjoni awtomatika ta' traduzzjonijiet u tħejji bażi empirika għat-TA. META-RESEARCH qed taħdem ma' oqsma u dixxiplini oħra, bħal tagħlim bil-magni u l-Komunità Semantika tal-Web. META-RESEARCH tiffoka fuq ġbir ta' dejta, tħejjija ta' settijiet ta' dejta u organizzazzjoni ta' riżorsi lingwistiċi għal skopijiet ta' evalwazzjoni; kumpilazzjoni ta' inventarji ta' għodod u metodi, u organizzazzjoni ta' workshops u avvenimenti ta' taħriġ għal membri tal-komunità. Din l-attività diġà identifikat b'mod ċar l-aspetti tat-TA fejn is-semantika tista' tħalli impatt fuq l-aħjar prassi kurrenti. Barra minn hekk, l-attività ħolqot rakkomandazzjonijiet dwar kif għandha tiġi ffaċċjata l-problema tal-integrazzjoni ta' informazzjoni semantika fit-TA. META-RESEARCH qiegħda wkoll tiffinalizza riżors ġdid lingwistiku għat-TA, il-Kampjun Ibridu Annotat tal-Korpus tat-TA, li jipprovdi dejta għall-pari ta' lingwi Ingliż-Ġermaniż, Ingliż-Spanjol u Ingliż-Ċek. META-RESEARCH żviluppat ukoll software li jiġbor korpora multilingwi li huma moħbija fuq il-web.

## Organizzazzjonijiet Membri

It-tabella li ġejja telenka l-organizzazzjonijiet u r-rappreżentanti tagħhom li jipparteċipaw fil-META-NET.

| Il-Pajjiż | L-Organizzazzjoni | Il-Parteċipant(i) |
|---|---|---|
| L-Awstrija | L-Università ta' Vjenna | Gerhard Budin |
| Il-Belġju | L-Università ta' Antwerp | Walter Daelemans |
| | L-Università ta' Leuven | Dirk van Compernolle |
| Il-Bulgarija | L-Akkademja tax-Xjenzi Bulgarija | Svetla Koeva |
| Il-Kroazja | L-Università ta' Zagreb | Marko Tadić |
| Ċipru | L-Università ta' Ċipru | Jack Burston |
| Ir-Repubblika Ċeka | L-Università Charles fi Praga | Jan Hajic |
| Id-Danimarka | L-Università ta' Copenhagen | Bolette Sandford Pedersen u Bente Maegaard |
| L-Estonja | L-Università ta' Tartu | Tiit Roosmaa |
| Il-Finlandja | L-Università ta' Aalto | Timo Honkela |
| | L-Università ta' Helsinki | Kimmo Koskenniemi u Krister Linden |

| Il-Pajjiż | L-Organizzazzjoni | Il-Parteċipant(i) |
|---|---|---|
| Franza | CNRS/LIMSI | Joseph Mariani |
| | L-Aġenzija tal-Evalwazzjonijiet u d-Distribuzzjoni tar-Riżorsi Lingwistiċi | Khalid Choukri |
| Il-Ġermanja | DFKI | Hans Uszkoreit u Georg Rehm |
| | L-Università ta' RWTH Aachen | Hermann Ney |
| | L-Università ta' Saarland | Manfred Pinkal |
| Il-Greċja | L-Istitut għall-Ipproċessar tal-Lingwa u t-Taħdit, "Athena" R.C. | Stelios Piperidis |
| L-Ungerija | L-Akkademja tax-Xjenzi l-Ungerija | Tamás Váradi |
| | L-Università tat-Teknoloġija u l-Ekonomija Budapest | Géza Németh u Gábor Olaszy |
| L-Islanda | L-Università tal-Islanda | Eirikur Rögnvaldsson |
| L-Irlanda | L-Università tal-Belt ta' Dublin | Josef van Genabith |
| L-Italja | Kunsill Nazzjonali tar-Riċerka, Istitut tal-Lingwistika Kompjutazzjonali "Antonio Zampolli" | Nicoletta Calzolari |
| | Fundazzjoni Bruno Kessler | Bernardo Magnini |
| Il-Latvja | Tilde | Andrejs Vasiljevs |
| | L-Istitut tal-Matematika u x-Xjenza tal-Kompjuter, L-Università tal-Latvja | Inguna Skadina |
| Il-Litwanja | L-Istitut tal-Lingwa Litwana | Jolanta Zabarskaitė |
| Il-Lussemburgu | Arax Ltd. | Vartkes Goetcherian |
| Malta | L-Università ta' Malta | Mike Rosner |
| Il-Pajjiżi l-Baxxi | L-Università ta' Utrecht | Jan Odijk |
| | L-Università ta' Groningen | Gertjan van Noord |
| In-Norveġja | L-Università ta' Bergen | Koenraad De Smedt |
| Il-Polonja | L-Akkademja Pollakka tax-Xjenzi | Adam Przepiórkowski u Maciej Ogrodniczuk |

| Il-Pajjiż | L-Organizzazzjoni | Il-Parteċipant(i) |
|---|---|---|
| | L-Università ta' Lodz | Barbara Lewandowska-Tomaszczyk and Piotr Pęzik |
| Il-Portugall | L-Università ta' Lisbona | Antonio Branco |
| | L-Istitut tas-Sistemi tal-Inġinerija u l-Kompjuters | Isabel Trancoso |
| Ir-Rumanija | L-Akkademja tax-Xjenzi Rumanija | Dan Tufis |
| | L-Università ta' Alexandru Ioan Cuza | Dan Cristea |
| Is-Serbja | L-Università ta' Belgrad | Dusko Vitas, Cvetana Krstev u Ivan Obradovic |
| | L-Istitut ta' Mihailo Pupin | Sanja Vranes |
| Is-Slovakkja | L-Akkademja tax-Xjenzi Slovakkja | Radovan Garabik |
| Is-Slovenja | L-Istitut Jozef Stefan | Marko Grobelnik |
| Spanja | Barcelona Media | Toni Badia |
| | L-Università Teknika ta' Catalonia | Asunción Moreno |
| | L-Università ta' Pompeu Fabra | Núria Bel |
| L-Isvezja | L-Università ta' Gothenburg | Lars Borin |
| Ir-Renju Unit | L-Università ta' Manchester | Sophia Ananiadou |
| | L-Università ta' Edinburgu | Steve Renals |

# Referenzi

---

i European Commission Directorate-General Information Society and Media, *User language preferences online*, Flash Eurobarometer #313, 2011 (http://ec.europa.eu/public_opinion/flash/fl_313_en.pdf).

ii European Commission, *Multilingualism: an asset for Europe and a shared commitment*, Brussels, 2008 (http://ec.europa.eu/education/languages/pdf/com/2008_0566_en.pdf).

iii UNESCO Director-General, *Intersectoral mid-term strategy on languages and multilingualism*, Pariġi, 2007 (http://unesdoc.unesco.org/images/0015/001503/150335e.pdf).

iv European Commission Directorate-General for Translation, *Size of the language industry in the EU*, Kingston Upon Thames, 2009 (http://ec.europa.eu/dgs/translation/publications/studies).

v L-awturi huma obbligati lejn il-Professur Ray Fabri, li l-artiklu li jmiss, iċċitat fil-biblijografija bħala Fabri (dalwaqt jiġi ppubblikat[a]), kien l-ispirazzjoni għal ħafna mill-kontenut u ħafna mill-eżempji f'din it-taqsima.

vi http://www.nso.gov.mt/statdoc/document_file.aspx?id=2840

vii Għal rapport bir-reqqa dwar l-istorja lingwistika ta' Malta ara Brincat (2011).

viii *l-Orizzont* mis-7 ta' Settembru, 1995; riprodott f' Ambros 1998: p. 135 l-Orizzont mis-7 ta' Settembru, 1995; riprodott f'Ambros 1998: p. 135

ix Għal rapport sħiħ dwar tempi u aspetti fil-Malti, ara Fabri (1995) u Ebert (2000)

x http://www.kunsilltalmalti.gov.mt/

xi http://www.akkademjatalmalti.com/

xii http://www.fb10.uni-bremen.de/ghilm/default.aspx

xiii http://www.nso.gov.mt/statdoc/document_file.aspx?id=2651

xiv http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/556&format=HTML&aged=0&language=EN&guiLanguage=en

xv "Maltese language hardly used on the internet", *The Times of Malta*, 16/05/2011.

xvi L-awturi huma obbligati lejn Dr Albert Gatt (L-Istitut tal-Lingwistika, L-Università ta' Malta) talli ġibed l-attenzjoni tagħhom għal dan id-dokument.

xvii http://www.gov.mt/

xviii http://www.nettv.com.mt/

xix http://www.one.com.mt

xx http://www.rtk.org.mt/

xxi http://www.pbs.com.mt/

xxii http://www.radio101.com.mt/

xxiii http://eur-lex.europa.eu

xxiv http://langtech.jrc.it/JRC-Acquis.html

xxv http://mlrs.research.um.edu.mt/

xxvi see http://linux.org.mt/spellcheck

xxvii http://www.spiegel.de/netzwelt/web/0,1518,619398,00.html

xxviii http://www.pcworld.com/businesscenter/article/161869/google_rolls_out_semantic_search_capabilities.html

xxix see http://www.philb.com/cse/malta.htm

xxx see http://www.charonite .com

xxxi http://www.let.uu.nl/lt4el/

xxxii Aktar ma jkun għoli l-punteġġ, aktar tkun tajba t-traduzzjoni. Traduttur uman iġib madwar 80. K. Papineni, S. Roukos, T. Ward, W.-J. Zhu. BLEU: A Method for Automatic Evaluation of Machine Translation. In Proceedings of the 40th Annual Meeting of ACL, Philadelphia, PA.

xxxiii ingħata isem mill-ġdid "Department of Intelligent Computer Systems (ICS)" fl-2009

xxxiv (Galea (1999), Mangion (1999), Farrugia (1999), Farrugia (2000), Mizzi (2000), Bajada (2004), Attard(2005), Farrugia (2008), Farrugia (2009), Vella (2010)

xxxv M. Rosner, R. Fabri, J. Caruana, M. Montebello u oħrajn

xxxvi ILSIENNA/Our Language: Journal of the International Association of Maltese Linguistics (GHILM), Universitätsverlag Brockmeyer, Bochum.

xxxvii Crimson Wing Ltd

xxxviii FITA (Fundazzjoni għall-Aċċess tat-TI)

xxxix A. Gatt, C. Borg, R. Fabri

xl M. Rosner