

The Human Language Model

Lane Schwartz

University of Illinois at Urbana-Champaign
lanes@illinois.edu

22 October 2014

The Human Language Model

A Re-examination of the Proper Role of Humans & Machines in
Interactive Language Translation 34 years after Kay (1980)

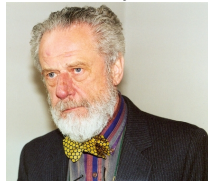
Lane Schwartz

University of Illinois at Urbana-Champaign
lanes@illinois.edu

22 October 2014

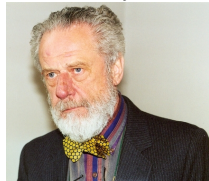
“ *The world is badly in need of translators. Almost nobody denies this. The number of pairs of languages between which translations must be made and the number and types of documents involved is constantly increasing.* ”

Martin Kay (1980)



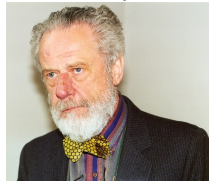
“ *But we are fortunate to be children of the age of computers and it is to them that we naturally turn. A computer is a device that can be used to magnify human productivity.* ”

Martin Kay (1980)



“ Properly used, it does not dehumanize by imposing its own Orwellian stamp on the products of the human spirit and the dignity of human labor but, by taking over what is mechanical and routine, it frees human beings for what is essentially human. ”

Martin Kay (1980)



The Human Translation Model



CAT Tools



Bilingual Interactive Translation



Monolingual Post-Editing

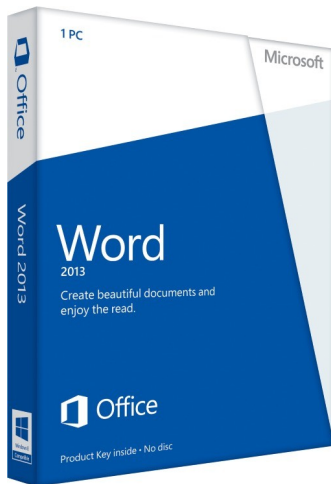


A Way Forward



References

The Human Translation Model



CAT Tools
●○○○○○○○○○○○○○○○○○○○○

Bilingual Interactive Translation
○○○○○○○

Monolingual Post-Editing
○○○○○○○

A Way Forward
○○

References

CAT Tools



Bilingual Interactive Translation



Monolingual Post-Editing



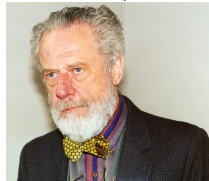
A Way Forward



References

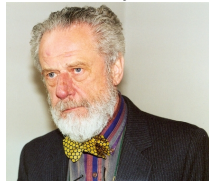
“ *There was a long period — for all I know, it is not yet over — in which the following comedy was acted out nightly in the bowels of an American government office with the aim of rendering foreign texts into English.* ”

Martin Kay (1980)



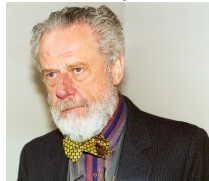
“ *Passages of innocent prose on which it was desired to effect this delicate and complex operation were subjected to a process of vivisection at the hands of an uncomprehending electronic monster that transformed them into stammering streams of verbal wreckage.* ”

Martin Kay (1980)



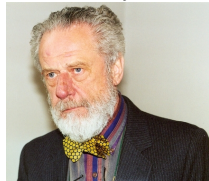
“ *Simple tools that would have done so much to make the repair work easier and more effective were not to be had ...* ”

Martin Kay (1980)

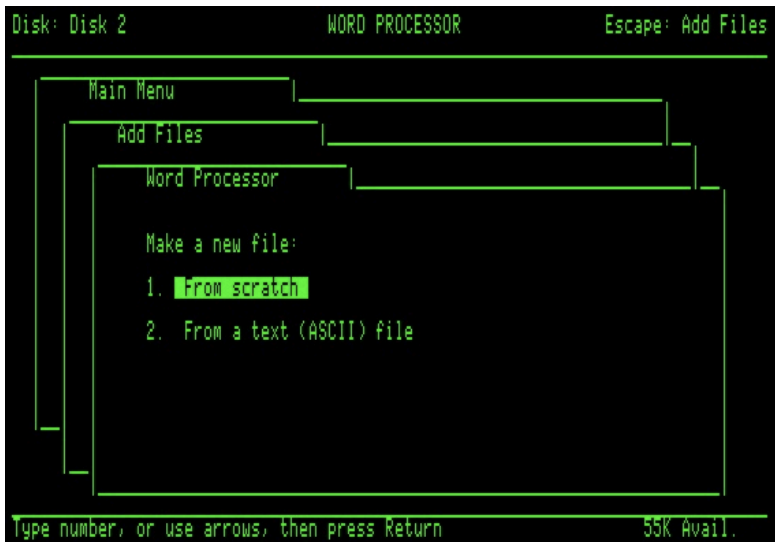


“ ...we can surely take a moment to wonder at the follies of the past and nostalgically to muse about what a kinder and more rational world would be like.”

Martin Kay (1980)



1980s word processing



Xerox PARC



CAT Tools



Bilingual Interactive Translation



Monolingual Post-Editing



A Way Forward



References

READY: Select operand or type command
Last command was LOOK
[A_substa...!_way_] {Computer...XEROX_]S

Personal Distributed Computing The Alto and Ethernet Software

*Butler W. Lampson
Digital Equipment Corp. Systems Research Center*

Abstract

The personal distributed computing system based on the Alto and the Ethernet was a major effort to make computers help people to think and communicate. A complex and diverse collection of software was built to pursue this goal, ranging from operating systems, programming environments, and communications software to printing and file servers, user interfaces, and applications such as editors, illustrators, and mail systems.

1. Introduction

A substantial computing system based on the Alto [Thacker et al.

Computer Science Laboratory
Xerox Palo Alto Research Center
3333 Goyate Hill Road
Palo Alto, California 94304

XEROX

Glen J. Culler
608 Lichfield Lane
Santa Barbara, CA 93109

Dear Glen:

This is a follow-up to earlier correspondence you received from Alan Perlis regarding the ACM Conference on the History of Personal Workstations. As you know, the conference is scheduled for January

CAT Tools

Chapter III

III - 1 - BUT DE L'ANALYSE SYNTAXIQUE

Dans le sens indiqué par V. YNGVE, l'analyse syntaxique associée à chaque phrase du texte un spécificateur structurale. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition.

Il s'agit d'une part de la présentation sous forme de constituents immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituents versus" a été l'un des sujets débattus au cours de

READY: Select operand or type command
Last command was LOOK
[A_substa...!_way_] {Computer...XEROX_]S

Personal Distributed Computing The Alto and Ethernet Software

*Butler W. Lampson
Digital Equipment Corp. Systems Research Center*

Abstract

The personal distributed computing system based on the Alto and the Ethernet was a major effort to make computers help people to think and communicate. A complex and diverse collection of software was built to pursue this goal, ranging from operating systems, programming environments, and communications software to printing and file servers, user interfaces, and applications such as editors, illustrators, and mail systems.

1. Introduction

A substantial computing system based on the Alto [Thacker et al.

Computer Science Laboratory
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, California 94304

XEROX

Glen J. Culler
608 Litchfield Lane
Santa Barbara, CA 93109

Dear Glen:

This is a follow-up to earlier correspondence you received from Alan Perlis regarding the ACM Conference on the History of Personal Workstations. As you know, the conference is scheduled for January

Chapter III

III - 1 - BUT DE L'ANALYSE SYNTAXIQUE

Dans le sens indiqué par V. YNGVE, l'analyse syntaxique associée à chaque phrase du texte un spécificateur structurale. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition.

Il s'agit d'une part de la présentation sous forme de constituents immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituents versus" a été l'un des sujets débattus au cours de

Chapter III

III - I - BUT DE L'ANALYSE SYNTACTIQUE

Dans le sens indiqué par V.YNGVE, l'analyse syntaxique associée à chaque phrase du texte un spécificateur structurale. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition.

Il s'agit d'une part de la présentation sous forme de constituents immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus" a été l'un des sujets débattus au cours de

III - I - AIMS OF SYNTACTIC ANALYSIS

In the sense indicated by V.YNGVE, syntactic analysis associates with each sentence of a text un spécificateur structurale. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition. Il s'agit d'une part de la présentation sous forme de constituents immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus" a été l'un des

Chapter III

III - 1 - BUT DE L'ANALYSE SYNTAXIQUE

Dans le sens indiqué par V.YNGVE, l'analyse syntaxique associe à chaque phrase du texte un spécificateur structurel. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition.

Il s'agit d'une part de la présentation sous forme de constituents immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus *structure*" a été l'un des sujets débattus au cours de

Spécificateur

Spécifier

Linguistique -
 Syntax - ●
 Description: *Phrase-marker (MK)*
 Sémantique - ●
 Spécifier

III - 1 - AIMS

In the sense of V.YNGVE, syntactic analysis associates with each sentence of a text a structural specifier. In 1963, epoch to which the C. E. T. A. began to occupy itself with syntax, this specifier is represented by a tree structure and corresponds to the linguistic level called "surface syntax". In what concerns the presentation of these tree structures, two schools, founded on two different linguistic conceptions, confront each other and enter into competition. It is a question of the presentation under the form of immediate constituents and of the other part of the presentation under the form of dependencies. Even though H. GAIFMAN has shown in 1965 that any language describable in one of these two systems is also describable in the other [14], it must be believed that the philosophical problem remains since the theme "constituency versus *structure*" has been one of the subjects debated during the course of

Chapter III

III - 1 - BUT DE L'ANALYSE SYNTACTIQUE

Dans le sens indiqué par V.YNGVE, l'analyse syntaxique associe à chaque phrase du texte un spécificateur structurel. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition.

Il s'agit d'une part de la présentation sous forme de constituants immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus dependency" a été l'un des sujets

III - 1 - AIMS OF SYNTACTIC ANALYSIS

In the sense of V.YNGVE, syntactic analysis associates with each sentence of a text a [2 structural description]. In 1963, the time when C. E. T. A. started concerning itself with syntax, this [2 structural description] took the form of a tree and corresponded to the linguistic level called "surface structure". As for the [presentation] of these trees, two schools, based on two different linguistic conceptions, confronted one another and began to compete. On the one hand there was the [presentation] in the form of immediate constituents and, on the other, the [presentation] in the form of [dependency]s. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus" a été l'un des sujets débattus au cours de

Chapter III

III - 1 - BUT DE L'ANALYSE SYNTACTIQUE

Dans le sens indiqué par V.YNGVE, l'analyse syntaxique associe à chaque phrase du texte un spécificateur structurel. En 1963, époque à laquelle le C. E. T. A. a commencé à s'occuper de syntaxe, ce spécificateur est représenté par une arborescence et correspond au niveau linguistique appelé "syntaxe de surface". En ce qui concerne la présentation de ces arborescences, deux écoles, fondées sur deux conceptions linguistiques différentes, s'affrontent et entrent en compétition.

Il s'agit d'une part de la présentation sous forme de constituents immédiats et d'autre part de la présentation sous forme de dépendances. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus dependency" a été l'un des sujets

III - 1 - AIMS OF SYNTACTIC ANALYSIS

In the sense of V.YNGVE, syntactic analysis associates with each sentence of a text a [2 structural description]. In 1963, the time when C. E. T. A. started concerning itself with syntax, this [2 structural description] took the form of a tree and corresponded to the linguistic level called "surface structure". As for the [presentation] of these trees, two schools, based on two different linguistic conceptions, confronted one another and began to compete. On the one hand there was the [presentation] in the form of immediate constituents and, on the other, the [presentation] in the form of [dependency]s. Bien que H. GAIFMAN ait montré en 1965 que tout langage descriptible dans l'un quelconque de ces deux systèmes l'est aussi dans l'autre [14], il faut croire que le problème philosophique demeure puisque le thème "constituency versus" a été l'un des sujets débats au cours de

Computer Aided Translation



CAT Tools: Machine Translation

The screenshot displays the Sunda Translator interface. The main window shows a translation of a text about the Public Library of Science (PLoS). The source text is in English, and the target text is in Finnish. A term recognition window is open on the right, showing a message: "No open tembase." The interface includes a menu bar with options like "Project Settings", "Concordance Search", "Comments", "Messages", "Term Recognition", and "Termbase Search". The status bar at the bottom shows "All segments INS 69.27% 30.73% 0.00% Chars: 225".

Source Text (English):

The Public Library of Science (PLoS) is a non-profit organization of scientists and physicians committed to making the world's scientific and medical literature a freely available public resource.

The internet and electronic publishing enable the creation of public libraries of science containing the full text and data of any published research article, available free of charge to anyone, anywhere in the world.

Immediate unrestricted access to scientific ideas, methods, results, and conclusions will speed the progress of science and medicine, and will more directly bring the benefits of research to the public.

To realize this potential, a new business model for scientific publishing is required that treats the costs of publication as the final integral step of the funding of a research project.

To demonstrate that this publishing model will be successful for the publication of the very best research, PLoS will publish its own journals.

Target Text (Finnish):

Julkinen tiedekirjasto (PLoS) on tiedemiesten ja lääkäreiden yleishyödyllinen yhtiö, jotka ovat sitoutuneet tekemään maailman tieteellisestä ja lääketieteellisestä kirjallisuudesta vapaasti saatavissa olevan julkisen resurssin.

Internet ja sähköinen julkaisu toiminta mahdollistavat julkisten tiedekirjastojen homien, jotka sisältävät minkä tahansa julkaistun tutkimusjulkaisun täyttää tekstiä ja dataa, ilmaiseksi saatavissa, jokaiselle, kaikkialla maailmassa.

Term Recognition Window:

No open tembase.

2 **Finding a location for your photo printer**

3 Place the photo printer on a flat, clean and dust-free surface,
of direct sunlight **pt11**.

4 Allow at least 12 cm clearance from the back of the photo printer.

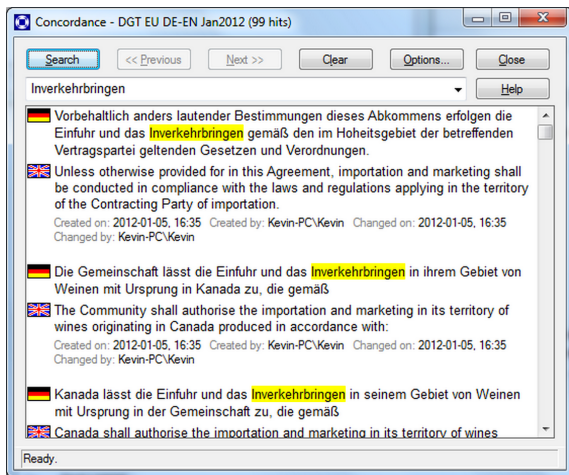
5 When connecting power or USB cables, keep the cables clear
photo printer.

6 For proper ventilation **ph14**, make sure the top and back of the

printer ✕

One who makes prints.
The operator of a printing press or owner of a printing business.
(context: computing: lang=en) A device, usually attached to a computer, used to print text or images onto paper. Also see plotter.

[More>>](#)



CAT Tools: Terminology Management

SamplePhotoPrinter.doc		SamplePhotoPrinter.doc	
1	Getting Started	OK	Erste Schritte
2	Finding a location for your photo printer	91%	Geeigneten Aufstellungsort für Ihren Fotodrucker finden
3	Place the photo printer on a flat, clean and dust-free surface, in a dry location, and out of direct sunlight. Allow at least 12 cm clearance from		

CAT Tools
○○○○○○○○○○○○○○○○○○○○

Bilingual Interactive Translation
○○○○○○○

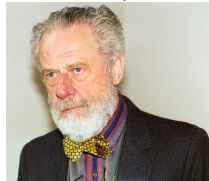
Monolingual Post-Editing
○○○○○○○

A Way Forward
○○

References

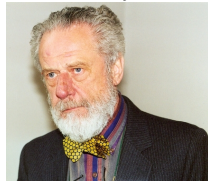
“ I want to advocate a view of the problem in which machines are gradually, almost imperceptibly, allowed to take over certain functions in the overall translation process. ”

Martin Kay (1980)



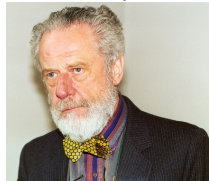
“ *First they will take over functions not essentially related to translation. Then, little by little, they will approach translation itself.* ”

Martin Kay (1980)



“ *The keynote will be modesty. At each stage, we will do only what we know we can do reliably. Little steps for little feet!* ”

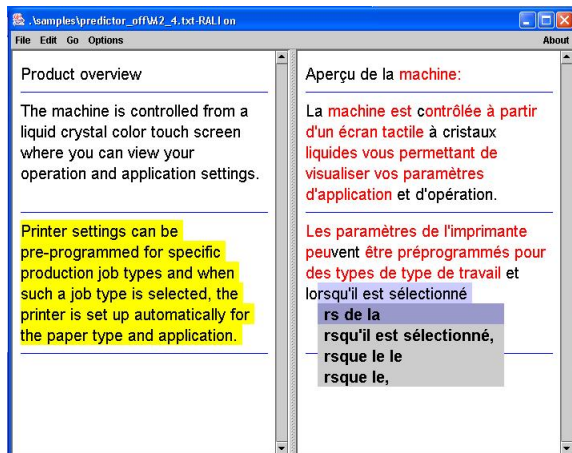
Martin Kay (1980)



Bilingual Interactive Translation

Pop-up typing completion

TransType (Langlais et al., 2000; Barrachina et al., 2009)



Visualization of Translation Options and Edit Distance

Caitra (Koehn, 2009)

[1] Sarkozy à la rencontre des pêcheurs en colère [X]
[2] Avant de s'envoler pour les Etats-Unis, Nicolas Sarkozy va faire un détour par la Bretagne. [3] Face à la démonstration de force des pêcheurs qui ont fortement perturbé lundi les accès de plusieurs ports bretons et normands, le président de la République a décidé de se rendre mardi matin au port du Guilvinec, dans le Finistère, pour y rencontrer leurs représentants. [4] Ces derniers ont promis de l'accueillir "dans un esprit constructif" et espèrent qu'il "viendra avec des propositions concrètes".

[1] Sarkozy à la rencontre des pêcheurs en colère >>

Sarkozy at the meeting of fishermen to angry meet (8 edits)

Sarkozy to meet

enter save [X]

Sarkozy	à	la	rencontre	des	pêcheurs	en	colère	
Sarkozy	the	meeting	the	fishermen	angry			
	in	the	meeting	the	fishermen	in	anger	
		to	the				anger	
				fishermen			angry	
			the	meeting	of	the	fishermen	
			the	meeting	of		as	
			the	meeting			by	
			the		of	in	the	anger ,
			to			at	however ,	
			at			by	however	

Visualization of Confidence Measures

Casmatcat (Alabau et al., 2014)

With a cholera epidemic raging, and more than 1m earthquake survivors still living in tents, there were fears that turnout would be low.

Con una epidemia de cólera virulentos, y más de 1 millón supervivientes del terremoto que siguen viviendo en tiendas de campaña, existen temores de que la participación será baja.

ITP T- DRAFT TRANSLATED

Translation matches

With a cholera epidemic raging, and more than 1m earthquake survivors still living in tents, there were fears that turnout would be low.

Con una epidemia de cólera virulentos, y más de 1 millón supervivientes del terremoto que siguen viviendo en tiendas de campaña, existen temores de que la participación será baja.

Source: ITP Thu Mar 07 2013 14:00:43 GMT+0300 (CET) 56

Source match Target match Replacement Case sensitive Regular expression Replace View Rules

Progress: 0% Total Words: 281 To-do: 281 Speed: --- Words/h Completed in: --- Reset Document

Rating Segment Confidence

TransCenter (Denkowski et al., 2014)

Talk 1 (practice)			Pause	Submit	?
	Source	Translation	Rating		
1	Muchas gracias Chris. Y es en verdad un gran honor	Thank you so much, Chris. And it's truly a great honor	5 - Very Good ▾		
2	tener la oportunidad de venir a este escenario por segunda vez. Estoy extremadamente agradecido.	to have the opportunity to come to this stage twice. I'm extremely grateful	Rate Translation ▾		

Vertically interleaved layout

Green et al. (2014a,b)

À équiper le centre de formation Studeo qui est accessible aux personnes à mobilité réduite et dont nous travaillons à la réalisation dans le cadre de l'institut Jedlička, avec l'association Tap, et ça depuis six ans.

To equip studeo training centre which is accessible to people with reduced mobility and we work to achieve in the framework of the Institute jedlička, with tap, and been there for six years.

Des enseignants se rendent régulièrement auprès des élèves de l'institut Jedličkův et leur proposent des activités qui les intéressent et les amusent.

Teachers regularly visit Jedličkův Institute students and offered them activities of interest to them and having fun.

Les étudiants eux-mêmes n'ont pas les moyens de se rendre à des cours, nous essayons de les aider de cette manière.

The students themselves cannot be required to attend courses, we are trying to help

Dans le cadre de l'Institut Jedlička, nous transférerons ce projet dans un no

themselves cannot be required
themselves cannot
themselves could not
themselves do not
themselves cannot afford

Translation completion with source lookup

Green et al. (2014a,b)

The screenshot shows a translation tool interface with three main sections:

- Source Text:** "To equip studeo trai... reduced mobility and... Institute jedlička,"
- Target Text:** "Des enseignants se rendent régulièrement auprès... proposent des activités qui les intéressent et les..."
- Completion Section:** "Teachers regularly visit Jedličkôv... activitie... them and... Les étudiant... s les moyens... aider de cett..."

A dropdown menu is open over the word "regularly" in the completion section, showing the following suggestions:

- routinely
- steadily
- regular
- regularly

CAT Tools
○○○○○○○○○○○○○○○○○○○○

Bilingual Interactive Translation
○○○○○○○

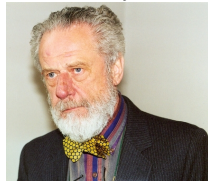
Monolingual Post-Editing
○○○○○○○

A Way Forward
○○

References

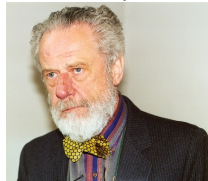
“ *To translate is to re-express in a second language what has been understood by reading a text.* ”

Martin Kay (1980)



“ *Any purported solution to the problem that does not involve understanding in some sense is, at best, ad hoc.* ”

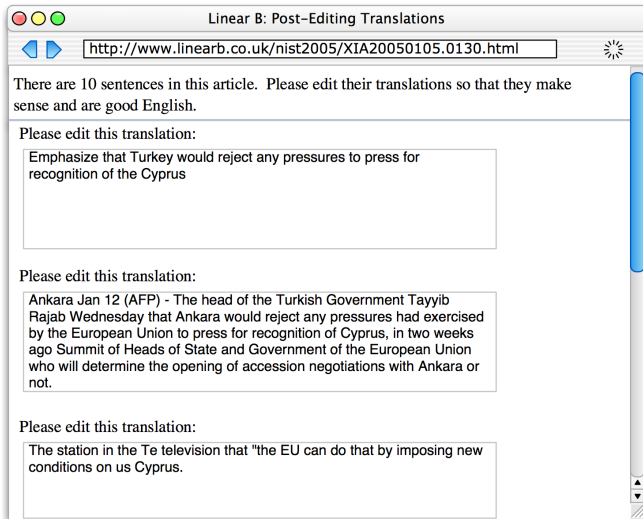
Martin Kay (1980)



Monolingual Post-Editing

Monolingual Post-Editing

(Callison-Burch, 2005)



Monolingual Post-Editing

(Mitchell et al., 2013)

Bearbeiten

Klicken Sie auf den Text um diesen zu bearbeiten:

Lösung für Windows XP nicht mehr Windows Updates von geschieft - ComboFix.exe -----
- Website www.bleepingcomputer.com/download/combofix/ und laden Sie ComboFix.exe ausführen. Sehr viel Geduld. Es geht aber wie 100 Phasen und Neustarts, aber es funktioniert. Ich habe für mich! ----- Benutzer NICHT combofix ausgebildet ausführen, ohne Hilfestellung, es ist ein leistungsstarkes Tool, das kann zu Problemen mit dem System Selbst die Schöpfer Staat und nicht, um sie auszuführen ich hatte, System wieder nach der Ausführung combofix und andere Tools, aber ich bin in den richtigen Bereich zu tun. Und es ist nicht zu 100 Phasen, Thread Ersteller kann nicht einmal das richtig.

Aktueller Satz zum Bearbeiten:

Lösung für Windows XP nicht mehr Windows Updates von geschieft - ComboFix.exe

Zurück Weiter

Kommentare ?

Aufgabe fertigstellen Richtlinien für später speichern

Monolingual Translation Options

(Callison-Burch, 2005)

Linear B: Visualization Tool

<http://www.linearb.co.uk/nist2005/segment1.html>

اردوغان	يؤكد	بأن	ترکيا	سترفض	اي	ضغوطات	لحثها	على	الاعتراف	بقبرص
		that turkey that turkey is turkey that turkey had that turkey will						to recognize recognition recognition of to recognize the the recognition		
	confirm that the confirm that affirm that confirms that who assert that our						urge it to to urge them to and urge it to order to urge it to urging			
ardogan ardogan ,	affirms stresses emphasize confirms emphasizes	that that the to , that that ,	turkey of turkey by turkey , turkey turkey ' s	reject will reject refuse not accept would be	any no of any i . e . meaning	pressure to pressure up to pressure	urge it urge and urge it urging urge them to	on to on the for to the	recognition recognition of recognized recognition of the	cyprus cyprus , to cyprus on cyprus cyprus and

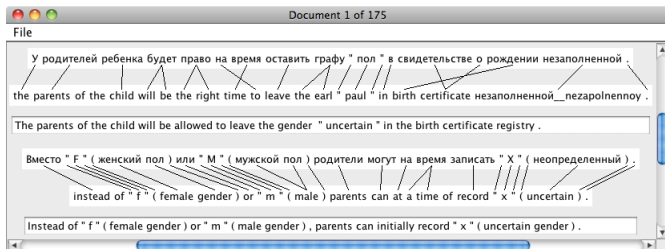
Monolingual Translation Options

(Koehn, 2010)

وكان	مجلس	التواب	الاميركي	استمد	الخميس	قانوننا	يطالب	يسحب	القوات	المقاتلة	الاميركية	من	العراق	في	موعد	اقصاء	الاول	من	تيسار	@ /	ايوزل
the	the us house of representatives	adopted	thursday	legally	calls for the withdrawal of	combat troops	us	iraq	in	no later than	the first	from	april								
the us house of representatives	the	thursday ,	law		the fighting forces	the us	from	iraq	in	the latest	the first of	april									
it was	the us house of representatives	adopted the	thu	the legally		the fighting forces	us	from	iraq	in		i	april								
he was	the us house of representatives	was adopted	thursday , the	the law	demands withdrawal of troops	fighter	the us		no later than	first	on	april									
he	the us house of representatives	was adopted	thursday 's	a law	calls for withdrawal of	combat forces		of	in the	not later than	first of										
he	the us house of representatives	was adopted	on thursday	a legally	calls for the withdrawal	forces	the fighter	from													
earlier ,		adopted a	on thursday ,	by law	demands the withdrawal of	troops		iraq													
was		was adopted	thursday the	legally	demands withdrawal of			of the													
it was the		adopted ,	thu ,	the legal	calls for withdrawal			from iraq in the													
earlier , the		adopted , the	thursday , a	legally @-@	demands the withdrawal			the american													
2008 ,	متحديا	مرة	جديدة	الرئيس	جورج	بوش	الذي	يعارض	اي	تحديد	موعد										
2008 ,	defying	once again	new	president george w. bush	which opposes	no date has been set for the															
the 2008	defiant	once again		president george bush	who opposes	no date has been set for															
2008	challenging	again	the new		which opposes	no date has been set															
	a defiant	the first			who opposes the	a date .															
	in defiance of	once again ,			who opposes	date .															
	challenging	once again the		president george bush , who	opposed to setting any	the date of the															
	in defiance	for the first time a new		president george w. bush 's	which opposes	no date															
in 2008 ,	defying the	again		us president george w. bush	opposed to	any															
	challenging the	time			who opposes the	date of															
	defying	once again , the			opposes	date															

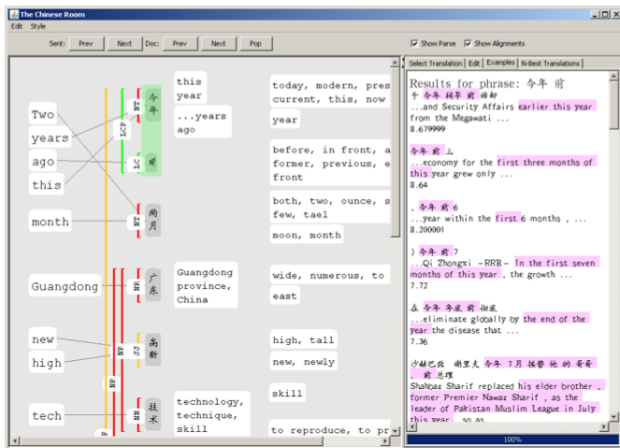
Monolingual Post-Editing with Alignments

(Schwartz et al., 2014)



Monolingual Post-Editing with Alignments

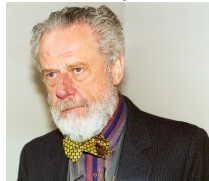
(Albrecht et al., 2009)



A Way Forward

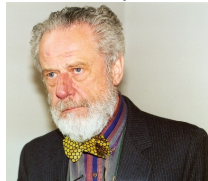
“ *The difference ...from a user's point of view is that here the translator has his say while the translation is underway whereas previously he had to wait.* ”

Martin Kay (1980)



“ *If this scheme can be made to work, as I claim it can, its many advantages are collectively overpowering.* ”

Martin Kay (1980)



- Commercial CAT Tools
- Research MT Systems
- Research Interactive MT
- Monolingual Translators

- **Commercial CAT Tools**

- Research MT Systems
- Research Interactive MT
- Monolingual Translators

- TM & MT integration
- Dictionary
- Concordance
- Terminology Management

- Commercial CAT Tools
 - **Research MT Systems**
 - Research Interactive MT
 - Monolingual Translators
- cdec
 - Joshua
 - Moses
 - Phrasal
 - ...

- Commercial CAT Tools
 - Research MT Systems
 - **Research Interactive MT**
 - Monolingual Translators
- Interactive typing completion
 - Partial translation MT extensions
 - Visualization of confidence measures

- Commercial CAT Tools
 - Research MT Systems
 - Research Interactive MT
 - **Monolingual Translators**
- Machine translation model
 - Human language model
 - Decoding shared between human and machine

- Vicent Alabau, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Ulrich Germann, Jesús González-Rubio, Robin Hill, Philipp Koehn, Luis Leiva, Bartolomé Mesa-Lao, Daniel Ortiz-Martínez, Herve Saint-Amand, Germán Sanchis Trilles, and Chara Tsoukala. Casmacat: A computer-assisted translation workbench. In *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 25–28, Gothenburg, Sweden, April 2014.
- Joshua S. Albrecht, Rebecca Hwa, and G. Elisabeta Marai. Correcting automatic translations through collaborations between MT and monolingual target-language users. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 60–68, Athens, Greece, March–April 2009.
- Sergio Barrachina, Oliver Bender, Francisco Casacuberta, Jorge Civera, Elsa Cubel, Shahram Khadivi, Antonio Lagarda, Hermann Ney, Jesús Tomás, Enrique Vidal, and Juan-Miguel Vilar. Statistical approaches to computer-assisted translation. *Computational Linguistics*, 35(1): 3–28, March 2009.
- Chris Callison-Burch. Linear B system description for the 2005 NIST MT evaluation exercise. In *Proceedings of the NIST 2005 Machine Translation Evaluation Workshop*, 2005.
- Michael Denkowski, Alon Lavie, Isabel Lacruz, and Chris Dyer. Real time adaptive machine translation for post-editing with cdec and transcenter. In *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*, pages 72–77, Gothenburg, Sweden, April 2014.

References (cont.)

- Spence Green, Jason Chuang, Jeffrey Heer, and Christopher D. Manning. Predictive translation memory: A mixed-initiative system for human language translation. In *Proceedings of the 27th ACM User Interface Software and Technology Symposium (UIST '14)*, Honolulu, Hawai'i, October 2014a.
- Spence Green, Sida Wang, Jason Chuang, Jeffrey Heer, and Christopher D. Manning. Human effort and machine learnability in computer aided translation. In *Proceedings of the 2014 Conference on Empirical Methods on Natural Language Processing (EMNLP '14)*, Doha, Qatar, October 2014b.
- Martin Kay. The proper place of men and machines in language translation. Technical Report CSL-80-11, Xerox PARC, Palo Alto, California, October 1980. Reprinted in *Machine Translation*, 12(1–2):3–23, 1997.
- Philipp Koehn. A web-based interactive computer aided translation tool. In *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*, pages 17–20, Suntec, Singapore, August 2009.
- Philipp Koehn. Enabling monolingual translators: Post-editing vs. options. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 537–545, Los Angeles, California, June 2010.
- Philippe Langlais, George Foster, and Guy Lapalme. TransType: A computer-aided translation typing system. In *Proceedings of the ANLP/NAACL 2000 Workshop on Embedded Machine Translation Systems*, pages 46–51, Seattle, Washington, May 2000.

Linda Mitchell, Johann Roturier, and Sharon O'Brien. Community-based post-editing of machine translation content: monolingual vs. bilingual. In *Proceedings of the 2nd Workshop on Post-editing Technology and Practice (WPTP-2)*, pages 35–43, Nice, France, September 2013. EAMT.

Lane Schwartz, Timothy Anderson, Jeremy Gwinnup, and Katherine M. Young. Machine translation and monolingual postediting: The AFRL WMT-14 system. In *Proceedings of the Ninth Workshop on Statistical Machine Translation*, pages 186–194, Baltimore, Maryland, June 2014. Association for Computational Linguistics.