

## TrAva – a tool for evaluating Machine Translation – pedagogical and research possibilities

Belinda Maia, Diana Santos, Luís Sarmento & Anabela Barreiro - Linguateca

## Why MT Matters

- Social reasons
- Political reasons
- Commercial importance
- Scientific and philosophical interest

## Useful bibliography

- ARNOLD, D, BALKAN, L., LEE HUMPHREYS, R. Lee, MEIJER, S., & SADLER, S. (1994) *Machine Translation - An Introductory guide*. Manchester & Oxford : NCC Blackwell. ISBN 1-85554-217-X - or at: <http://www.essex.ac.uk/linguistics/clmt/MTbook/>.
- COLE, Ron (ed) 1996 "Survey of the State of the Art in Human Language Technology" Chapter 8 - Multilinguality, at the Center for Spoken Language Understanding, Oregon. <http://cslu.cse.ogi.edu/HLTsurvey/ch8node2.html#Chapter8>
- MELBY, Alan K. 1995, *The Possibility of Language: A discussion of the Nature of Language, with implications for Human and Machine Translation*. Amsterdam: John Benjamins Pub. Co.

## Machine Translation (MT) – a few dates

- 1947 Warren Weaver - his ideas led to heavy investment in MT
- 1959 Bar-Hillel - considers FAHQMT - FULLY AUTOMATIC HIGH QUALITY MACHINE TRANSLATION philosophically impossible
- 1964 ALPAC Report - on limitations of MT > withdrawal of funds
- late 1970s the CEC purchase of SYSTRAN and beginning of EUROTRA project.
- Upward trend in the 1970s and 1980s
- Today: MT technology - high-end versus low-end systems
- MT and the Internet

From Arnold et al 1995

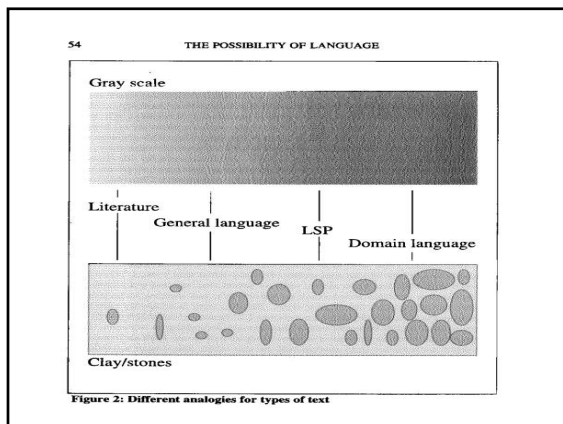
### Some Popular Misconceptions about MT

- **False:** MT is a waste of time because you will never make a machine that can translate Shakespeare.
- **False:** There was/is an MT system which translated *The spirit is willing, but the flesh is weak* into the Russian equivalent of *The vodka is good, but the steak is busy, and hydraulic runs into the French equivalent of wasser goat*. MT is useless.
- **False:** Generally, the quality of translation you can get from an MT system is very low. This makes them useless in practice.
- **False:** MT threatens the jobs of translators.
- **False:** The Japanese have developed a system that you can talk to on the phone. It translates what you say into Japanese, and translates the other speaker's replies into English.
- **False:** There is an amazing South American Indian language with a structure of such logical perfection that it solves the problem of designing MT systems.
- **False:** MT systems are machines, and buying an MT system should be very much like buying a car.

From Arnold et al 1995

### Some Facts about MT

- **True:** MT is useful. The METEO system has been in daily use since 1977. As of 1994, it was regularly translating around 45 000 words daily. In the 1980s, The diesel engine manufacturer Perkins Engines was saving around £2000 and up to 15 weeks on each manual translation.
- **True:** While MT systems sometimes produce howlers, there are many situations where the ability of MT systems to produce reliable, if less than perfect, translations at high speed is valuable.
- **True:** In some circumstances, MT systems can produce good quality output: less than 4% of METEO output requires any correction by human translators at all (and most of these are due to transcription errors in the original texts). Even where the quality is lower, it is often easier and cheaper to revise 'crack quality' MT output than to translate entirely by hand.
- **True:** MT does not threaten translators' jobs. The need for translation is vast and unlikely to diminish, and the limitations of current MT systems are too great. However, MT systems can take over some of the boring, repetitive translation jobs and allow human translators to concentrate on more interesting tasks, where their specialist skills are really needed.
- **True:** Speech-to-Speech MT is still a research topic. In general, there are many open research problems to be solved before MT systems will be close to the abilities of human translators.
- **True:** Not only are there are many open research problems in MT, but building an MT system is an arduous and time-consuming job, involving the construction of grammars and very large monolingual and bilingual dictionaries. There is no 'magic solution' to this.
- **True:** In practice, before an MT system becomes really useful, a user will typically have to invest a considerable amount of effort in customizing it.



## MT architectures – Arnold et al

- Direct architecture - simple grammatical rules + a large lexical and phrasal database
- Transfer architecture - more complex grammar with an underlying approach of transformational-generative theory + considerable research into comparative linguistics in the two languages involved
- Interlingua architecture - L1 > a 'neutral language' (real, artificial, logical, mathematical..) > L2

## Major Methods, Techniques and Approaches today

- Statistical vs. Linguistic MT
  - assimilation tasks: lower quality, broad domains – *statistical techniques predominate*
  - dissemination tasks: higher quality, limited domains – *symbolic techniques predominate*
  - communication tasks: medium quality, medium domain – *mixed techniques predominate*
- Rule-based vs. Example-based MT
- Transfer vs. Interlingual MT
- Multi-Engine MT
- Speech-to-Speech Translation

MLIM - Multilingual Information Management: Current levels and Future Abilities - report (1999) Chapter 4 at: <http://www-2.cs.cmu.edu/~ref/mlim/chapter4.html>

## MT – present & future uses

- 'Gist' translation
- Ephemeral texts with tolerant users
- Human aided MT
  - Domain specific
  - Linear sentence structure
  - Pre-edited text or 'controlled language'
  - Post-editing
- Improvement of MT – particularly for restricted domains and registers

## MT and the Human translator

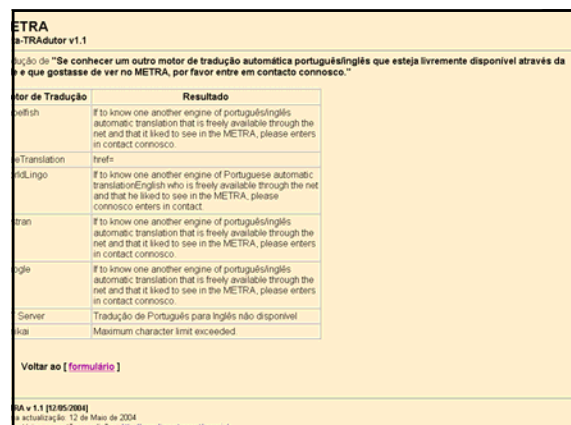
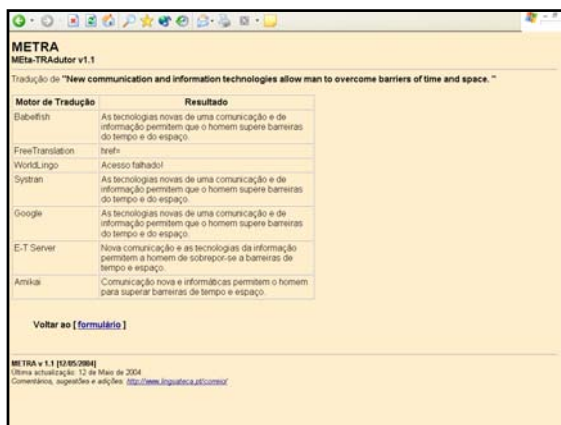
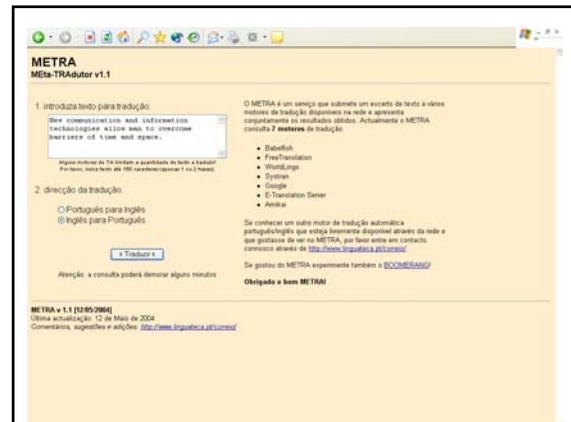
- MT is less of a threat to the professional human translator than English
- MT can encourage people's curiosity for texts in languages they do not understand > and lead to human translation
- MT can be a tool for the human translator
- Professional translators can learn to work with and train MT

## PoloCLUP's experiment

- Background
  - Master's seminar in Semantics and Syntax
  - Wish to raise students' awareness of the strengths and weaknesses of MT
  - Wish to develop their interest in MT as a tool
  - Need to improve their knowledge of linguistics.
  - Availability of free MT online
  - Automation of process provided by computer engineer

## Phase 1 - METRA

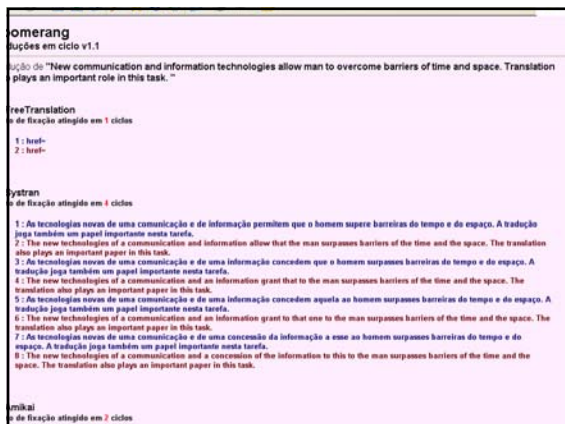
- <http://poloclup.linguateca.pt/ferramentas/metra/index.html>
- Translation using 7 online MT programmes
- EN > PT
- PT > EN
- At present this tool is getting about 60 hits per day!



## BOOMERANG

- <http://poloclup.linguateca.pt/ferramentas/boomerang/index.html>
- This tool submits a text for translation – and back-translation – and back-translation.... Until it reaches a fixed point
- This shows that the rules programmed for one language direction do not always correspond to the other language direction





## EVAL > TrAva

- Informal class experiment led to a useful research tool
- Several versions of EVAL
  - Different types of classification of input
  - Different explanations of errors of output
- Production, correction and re-correction of procedure interesting

## TrAva - procedure

- Online EN > PT MT using 4 MT systems:
  - Free Translation
  - Systran
  - E T Server
  - Amikai
- Researcher chooses area for analysis – e.g.
  - ambiguity
  - lexical and structural mismatches
  - Homographs and polysemous lexical items
  - syntactic complexity
  - multiword units: idioms and collocations
  - anaphora resolution

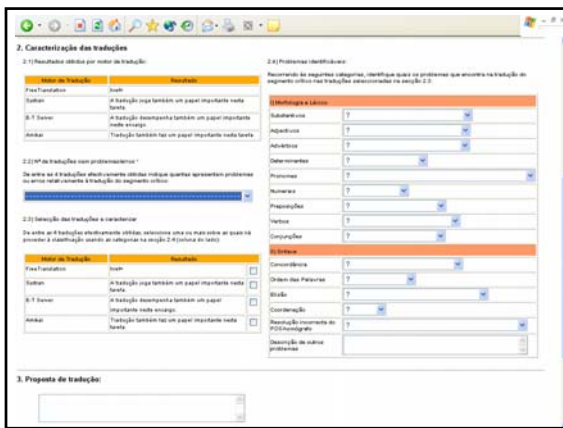
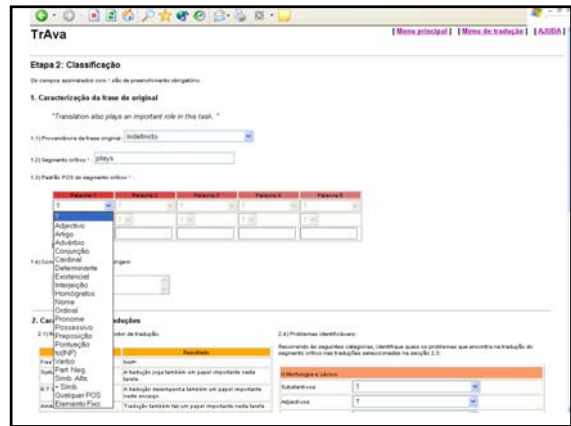
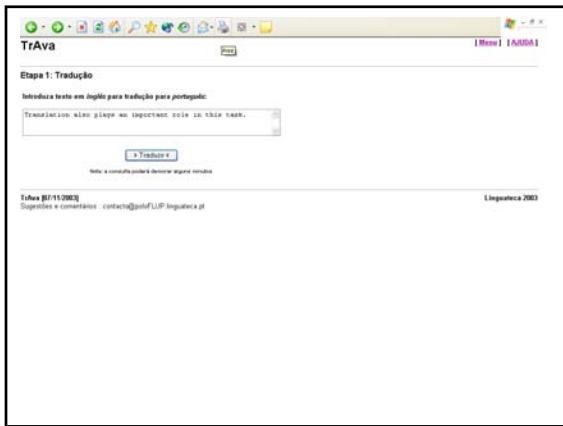
## TrAva - procedure

- Selection of **'genuine'** examples from BNC, Reuter's corpus, newspapers etc.
- Possible 'pruning' of unnecessary text (some systems accept limited text)
- No deliberate attempt to confuse the system
- BUT: avoidance of repetitive 'test suites'

## TrAva - procedure

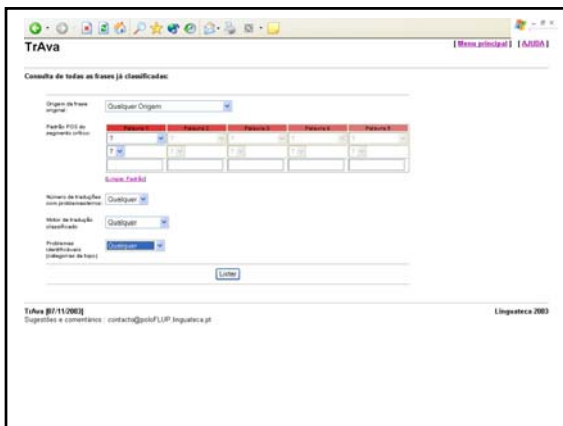
- Sentence submitted to TrAva
- MT results
- Researcher:
  - Classifies part of sentence being examined in terms of the English lexicon or POS (BNC codes)
  - Examines results
  - Explains errors in terms of Portuguese grammar





## Access to work done

- Researcher may access work done and review it
- Teacher / administrator can access student work and give advice
- FAQ



### Questões frequentes

- O que é Trava?**  
O Trava é uma ferramenta destinada a permitir a recolha colectiva de exemplos de tradução automática, do inglês para o português, classificados sistematicamente. Espera-se através desta recolha compilar um volume de informação categorizada suficientemente grande para poder observar padrões.
- Qualides que ocorrem no processo de tradução automática.**  
3. qualides que ocorrem no processo de tradução automática.
- Para que serve esta recolha?**  
Estes dados servem para melhorar o conhecimento sobre a tradução automática e servir de base a um corpus público de traduções automáticas. A recolha dos exemplos aplica também à avaliação contínua de ferramentas para futuras actividades de Avaliação de Tradução Automática (ATA) bem como para o estabelecimento de base-line.
- Que serviços de tradução automática é que o Trava consulta?**  
Actualmente o Trava consulta 2 motores de tradução:
  - FreeTranslation (<http://www.free-translation.com>)
  - System (<http://www.systeme.com/pt/ptan/ptan00>)
  - Tradução Google (<http://www.google.pt/translate/translate.html>)
  - Amk (<http://www.amk.com/pt/ptan/ptan00>)
- Porque foram escolhidos estes serviços de tradução e não outros?**  
O Trava recorre a estes serviços de tradução porque são todos diferentes. Há muitos outros serviços de tradução disponíveis na rede (Babelfish, Google, WordPerfect) mas que são baseados nos motores que foram seleccionados e por isso apresentam quase sempre os mesmos resultados. Além disso em muitos casos de resultados fustica e variada de classificação.
- Em que é que é baseado o sistema de classificação de POS usado para caracterizar o segmento crítico dos exemplos?**  
O sistema de classificação POS que utilizamos é o mesmo que foi empregado para realizar a análise POS do British National Corpus (BNC). A opção por este sistema de classificação deve-se essencialmente à grande disponibilidade do BNC nos meios linguísticos, o que torna na prática o seu sistema de análise POS um standard. Para mais informações acerca deste sistema de análise de POS, deve consultar o endereço <http://www.lingua.pt/pt/ptan/ptan00>.
- Trava usa métodos automáticos para fazer o que serve Trava?**  
Optar! Envie um e-mail para [contact@trava.lingua.pt](mailto:contact@trava.lingua.pt) (Óngued)

## Present situation

- METRA and BOOMERANG are all free to use online at:
- <http://poloclup.linguateca.pt/ferramentas>
- TraVa is free to use online at:
- <http://www.linguateca.pt/trava/>
- The corpus CoRTA – over 1000 sentences + 4 MT versions available for consultation at: <http://www.linguateca.pt/>



1	Endgore was big business <b>run</b> by ourselves	Info	
	PT: A vitória era muito grande de registar por aqueles		COMPARA
	ST: A vitória era registar grande facionando por aqueles		
	ET:		
	AM:		
2	An experiment <b>run</b> by Dr Ting is not a casual affair	Info	
	PT: Um		BNC
	ST: Uma experiência facionada por Dr. Ting não é um caso casual		
	ET:		
	AM: Uma experiência conta por Dr Ting não é um sfuero casual		
3	Para the depressed Helder and then set off on a spectacular 50-yard solo <b>run</b> down the right	Info	
	PT: Primeiro desgracia Helder e então pfo fora com omissão de 50 jardas espetacularmente corre para baixo à direita.		BNC
	ST: Primeiro desgracia Helder e quando se volta fora em um facionamento de solo de 50 jardas espetacularmente abaixo à direita.		
	ET:		
	AM: Primeiro ele desgracia Helder e então se parte em um 50 jardas espetacularmente corre de solo abaixo à direita.		
4	Discourse that will not <b>run</b> backward	Info	
	PT: Porque tempo não corre para trás		BNC
	ST: Porque tempo não facionando para trás		
	ET: Porque o tempo não facionando para trás		
	AM: Porque tempo não corre para trás		
5	Such <b>run</b> has been seen in the name of the people	Info	
	PT: Assim Almeida ficou confuso no nome do povo		?
	ST: Assim Domingo do facionando no nome dos povos		
	ET: Assim se Almeida ficou preocupado no nome do povo		
	AM:		

## Conclusions

- It has been a successful experiment
- It is useful pedagogically
  - As linguistic analysis
  - As appreciation of MT
- It has interesting theoretical implications > emphasis on 'real' sentences and recognition of interconnection of lexicon + syntax + context

## Conclusions

- Further work needs to be done on the classifications
  - E.g. the analysis of 'error' as 'lexical choice' needs to be able to combine with other possible reasons for error
- A lone researcher can use it to examine a restricted area
- BUT – a large team is needed to overhaul a system properly

## Homographs and Polysemy

- Homographs = words with same spelling but different syntactic use
- Polysemy = words with same spelling, but different meaning according to use or context
- BUT – the difference is not as clear-cut as all that
- However – major problem for MT



<b>10) frase Original:</b>	Richard has round brown eyes.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AB
significado FCIS:	[word="round" & pos="A/B"]
fraseTradução:	Richard tem olhos marrons redondos.
questões:	Richard tem em volta dos olhos marrons *
ET Server:	Acesso Falhado
resposta:	Richard tem redondos olhos marrons
número de traduções:	3
problemas identificados na tradução:	Adjetivos: Escolha Lexical
comentários sobre frase de origem:	Adaptada

<b>11) frase Original:</b>	When we round the next bend , bring your boat.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	VBB
significado FCIS:	[word="round" & pos="V/B"]
fraseTradução:	Quando nós arredondamos o lago dobrar, traga seu barco *
questões:	Quando nós arredondamos a curva seguinte, traga seu barco *
ET Server:	Acesso Falhado
resposta:	Quando nós arredondamos a próxima curva, traga seu barco *
número de traduções:	3
problemas identificados na tradução:	Verbs: Escolha Lexical
comentários sobre frase de origem:	Adaptada

<b>12) frase Original:</b>	I suppose you could make up a little compilation and send your friends round
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	ADV
fraseTradução:	Suponho-o pode compor uma compilação pequena e enviar seus amigos redondo *
questões:	Eu suponho que você poderia fazer alguma compilação pequena e enviar seus amigos redondos *
ET Server:	Suponho que você poderia compilar uma pequena compilação e enviar os seus amigos à volta de *
resposta:	Eu suponho-você poderia compor uma pequena compilação e poderia enviar seus amigos arredondados *
número de traduções:	4
problemas identificados na tradução:	Resolução: NVAGAD/Prep
comentários sobre frase de origem:	Talvez-SE para fazer uma pequena compilação e enviar os amigos para nem lá a casa. (Compara)
comentários sobre frase de destino:	Frase abreviada

<b>13) frase Original:</b>	One man had blood seeping through a bandage so he put round his head.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	ADV
fraseTradução:	Um homem teve sangue penetrar por uma compressa embebida em torno de a sua cabeça *
questões:	Um homem teve o sangue escorrer através de um bandage embebido em volta de sua cabeça *
ET Server:	Um homem teve sangue que gotejava através de uma bandagem embebida à volta da sua cabeça *
resposta:	Um homem teve sangue que vazou por uma bandagem embebida (circulo) a cabeça dela *
número de traduções:	3
problemas identificados na tradução:	Resolução: NVAGAD/Prep
comentários sobre frase de origem:	Havia um homem com sangue a escorrer de uma ligadura que tinha sido à cabeça
comentários sobre frase de destino:	Tradução-COMPARA

<b>14) frase Original:</b>	My name is Gloria and I run a beauty salon
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB VBB
significado FCIS:	
fraseTradução:	Chamo-me Gloria e eu como um salão de beleza *
questões:	Meu nome é Gloria e eu funciono um salão de beleza *
ET Server:	O meu nome é Gloria e eu executo um salão de beleza *
resposta:	Meu nome é Gloria e eu como um salão de beleza *
número de traduções:	4
problemas identificados na tradução:	Verbs: Escolha Lexical
comentários sobre frase de origem:	Assume por default "to run", "funcionar"

<b>15) frase Original:</b>	In the shirt run this new body it seen by the Commission as a useful means
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	Na camisa curta este novo corpo é visto pela Comissão como um meio útil *
questões:	No short funcionado este corpo novo é visto pelo comissão como um útil significa *
ET Server:	Acesso Falhado
resposta:	Na camisa curta este corpo novo é visto pela Comissão como uma meios útil *
número de traduções:	3
problemas identificados na tradução:	Substantivos: Escolha Lexical
comentários sobre frase de origem:	'Run' substantivo, precedido de um adjective 'short'. Será interessante ver até que ponto os motores de tradução identificam esta vertente
comentários sobre frase de destino:	Verbo escolhido: "run" como verbo

<b>16) frase Original:</b>	William Huskisson, President of the Board of Trade, was run down and killed at Parkside.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	Huskisson de william, Presidente da Tábua de Negócios, foi comido para baixo e matado em Parkside
questões:	William Huskisson, presidente da placa de comércio, foi funcionado para baixo e matado em Parkside
ET Server:	Acesso Falhado
resposta:	William Huskisson, Presidente da Tábua de Comércio, foi comido abaixo e matou a Parkside *
número de traduções:	3
problemas identificados na tradução:	Verbs: Escolha Lexical
comentários sobre frase de origem:	O motor não consegue traduzir 'was run down' no sentido de 'foi atropelado'. Assume por default 'foi comido abaixo'
comentários sobre frase de destino:	Tradução possível: "... foi atropelado e morto. "

<b>17) frase Original:</b>	She fell silent and a tear started to run down her cheek.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	Ela ficou silenciosa e uma lágrima começou a correr para baixo a bochecha dela
questões:	Cau silencioso e uma lágrima começou a correr para baixo a bochecha *
ET Server:	Acesso Falhado
resposta:	Ela se calou calada e uma lágrima começou a comer abaixo a bochecha dela
número de traduções:	3
problemas identificados na tradução:	Verbs: Escolha Lexical
comentários sobre frase de origem:	Apesar da tradução do Free Translation e do Amkar não estar muito correcta em termos de

<b>18) frase Original:</b>	lan never did the public displays of affection.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	O lan nunca gostou de exposições públicas de afecto *
questões:	lan nunca gostou de exposições públicas de afecto
ET Server:	lan nunca gostou de exposições públicas de afecto
resposta:	lan nunca gostou de exposições públicas de afecto
número de traduções:	3
problemas identificados na tradução:	Substantivos: Escolha Lexical
comentários sobre frase de origem:	O lan nunca gostou de exposições públicas de afecto.
comentários sobre frase de destino:	LOOC-2003, Examples bank
comentários sobre frase de destino:	O Amkar é o único motor que não traduz "displays" como "exposições"

<b>19) frase Original:</b>	The display changed suddenly from green to red.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	A exposição mudou repentinamente de verde a vermelho *
questões:	A exposição mudou de repente de verde ao vermelho.
ET Server:	A exposição passou de repente de verde a vermelho.
resposta:	A exibição mudou de repente de verde para vermelho.
número de traduções:	3
problemas identificados na tradução:	Substantivos: Escolha Lexical
comentários sobre frase de origem:	O motor mudou de repente de verde para vermelho.

<b>20) frase Original:</b>	I am attaching the relevant sections of the script.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	Estou colando as seções relevantes do texto *
questões:	Estou colando as seções pertinentes do texto *
ET Server:	Estou colando as seções pertinentes do texto *
resposta:	Estou colando as seções pertinentes do texto *
número de traduções:	3
problemas identificados na tradução:	Verbs: Escolha Lexical
comentários sobre frase de origem:	Estou em anexo as seções importantes do guião
comentários sobre frase de destino:	Utilização desnecessária do gerúndio

<b>21) frase Original:</b>	It was an impersonal object not in any way attached to her.
autor:	deberamos
origem:	7
palavra FCIS frase Original:	AVB
significado FCIS:	AVB
fraseTradução:	Era um objeto impessoal não em qualquer meio unido a ela *
questões:	Era um objeto impessoal não em toda a maneira unida a ela *
ET Server:	Não era um objeto impessoal em nenhuma maneira associada a ela *
resposta:	Não era de qualquer forma um objeto impessoal preso a ela *
número de traduções:	3
problemas identificados na tradução:	Verbs: Escolha Lexical
comentários sobre frase de origem:	Tratava-se de um objecto impessoal sem qualquer ligação com ela
comentários sobre frase de destino:	COMPARA, seguramente abreviada

<b>Frase Original:</b>	<b>The times of opening are Monday to Friday 14.00 - 16.00.</b>
palavra:	epítetos
origem:	OUTRA
PO3 Frase Original:	<b>NÃO CAC NMI</b>
palavra:	<b>times of opening</b>
origem:	times of opening
tradução:	As vezes de abertura são segunda-feira a sexta-feira 14:00 - 16:00. Os tempos de abertura são segunda-feira a sexta-feira 14:00 - 16:00. Os tempos de abrir são segunda-feira a sexta-feira 14:00 - 16:00. As vezes de abrir são segunda-feira a sexta-feira 14:00 - 16:00.
de tradução:	4
palavra identificada no texto:	Substantivos: Escolha Lexical Resolução: ang IVV
de proposta:	O horário de abertura é de segunda a sexta-feira, das 14:00 às 16:00 horas.
palavra sobre frase de:	LDCE 2003 (examples bank)
<b>Frase Original:</b>	<b>His question left an opening for me to say exactly what I thought.</b>
palavra:	epítetos
origem:	OUTRA
PO3 Frase Original:	<b>NMI</b>
palavra:	<b>opening</b>
origem:	opening
tradução:	Seu pergunta deixou uma abertura para mim para dizer exatamente o que eu pensava. His question left an opening for me to say exactly what I thought. A sua pergunta deixou uma abertura para mim de dizer exatamente o que eu pensava. A pergunta dele deixou uma abertura para eu dizer o que eu pensei exatamente.
de tradução:	4
palavra identificada no texto:	Substantivos: Escolha Lexical
de proposta:	A pergunta dele deixou a possibilidade de dizer exatamente o que eu pensei.

<b>Frase Original:</b>	<b>Gerardine did not watch the Golden Girls tonight because it clashed with News at Ten.</b>
palavra:	Racoonha
origem:	BNC
PO3 Frase Original:	<b>VVI ATB A3D NMI</b>
palavra:	<b>watch the Golden Girls</b>
origem:	O Gerardine não assistiu as Memórias Douradas hoje à noite porque chocou com notícias às dez. Gerardine não prestou atenção às memórias douradas hoje à noite porque clashed com notícias às dez. Gerardine não ficou as memórias de ouro hoje à noite porque esteve em conflito com notícias às dez. Gerardine não assistiu as Memórias Douradas hoje à noite porque colidiu com notícias às Dez.
de tradução:	3
palavra identificada no texto:	Verbo: Escolha Lexical
de proposta:	Gerardine não viu o programa "Golden Girls" hoje à noite porque coincidiu com "News at Ten". A motor não identificou que "Golden Girls" era um programa televisivo e, por isso, escolheu o significado "prestar atenção" e não "ver" para o verbo "watch".
<b>Frase Original:</b>	<b>They equipped I could watch the films in comfort.</b>
palavra:	Racoonha
origem:	BNC
PO3 Frase Original:	<b>PRP VMI VVI ATB NMI</b>
palavra:	<b>I could watch the films</b>
origem:	Assim equipado posso observar as películas em conforto. Assim equipado eu poderia prestar atenção às películas no conforto. Assim equipado poderia olhar as películas em conforto. Assim equipado eu poderia assistir os filmes em conforto.
de tradução:	3

<b>Frase Original:</b>	<b>Watch out for that yoghurt on top.</b>
palavra:	Racoonha
origem:	BNC
PO3 Frase Original:	<b>VVI AVB PRP DTB NMI</b>
palavra:	<b>Watch out for that yoghurt</b>
origem:	Tomar cuidado com aquele iogurte no topo. Retenjo para fora para esse iogurte no alto. Tenha cuidado a esse iogurte sobre topo. Assista em cima fora para aquele iogurte.
de tradução:	3
palavra identificada no texto:	Resolução: NVV
de proposta:	Cuidado com esse iogurte que está em cima!
palavra sobre frase de:	o verbo "watch" está no imperativo
palavra sobre a:	o motor traduziu ainda a partícula adverbial "out" e a preposição "for" separadas do verbo, por "para fora" e "para", respectivamente.
<b>Frase Original:</b>	<b>Behind you mister, watch out for the crocodile Cap'n Hook.</b>
palavra:	Racoonha
origem:	BNC
PO3 Frase Original:	<b>VVI AVB PRP ATB NMI</b>
palavra:	<b>watch out for the crocodile</b>
origem:	Atrás de você senhor, tome cuidado com o Gancho de crocodilo. Atrás de você o senhor, presta atenção para fora para o gancho de Cap'n do crocodilo. Atrás de você o senhor, tem cuidado ao crocodilo Cap'n Hook. Atrás de você o senhor, assista fora para o crocodilo Gancho de Cap'n.
de tradução:	2
palavra identificada no texto:	Verbo: Escolha Lexical

## Complex Noun Phrases

- DETERMINANTE + ADJECTIVO + NOME
- DETERMINANTE + ADJECTIVO COMPOSTO + ADJECTIVO + NOME
- DETERMINANTE + ADVÉRBIO (em -ly) + ADJECTIVO + ADJECTIVO + NOME

<b>Frase Original:</b>	<b>They have tiny translucent lance-shaped leaves and zebra-like flowers.</b>
palavra:	inútil
origem:	BNC
PO3 Frase Original:	<b>AD AD AD NMI</b>
palavra:	<b>translucent lance-shaped leaves and zebra-like flowers</b>
origem:	Têm tranlucido minúsculo lança formado sai de e assemelha-se zebra-like. Têm as folhas lance-dadas forma translúcidas minúsculas e assemblam-se a zebra-like. Eles têm folhas lança-amoldadas translúcidas minúsculas e se assemblam a hepáticas.
de tradução:	3
palavra identificada no texto:	Substantivos: Escolha Lexical Adjetivos: Escolha Lexical
<b>Frase Original:</b>	<b>These findings were confirmed by a joint British-Brazilian scientific team.</b>
palavra:	inútil
origem:	BNC
PO3 Frase Original:	<b>ATB AD AD AD NMI</b>
palavra:	<b>These findings were confirmed by a joint British-Brazilian scientific team</b>
origem:	Estes resultados foram confirmados por uma equipe científica, Britânica Brasileira em conjunto. Estes findings foram confirmados por uma equipe científica Britânico-british-Brazilian comum. Estes resultados eram confirmados por um time científico britânico-brasileiro em comum.
de tradução:	3
palavra identificada no texto:	Substantivos: Escolha Lexical Adjetivos: Escolha Lexical

<b>Frase Original:</b>	<b>Strategic vision cannot be achieved without a coherent sustained industrial strategy.</b>
palavra:	inútil
origem:	BNC
PO3 Frase Original:	<b>ATB AD AD AD AD NMI</b>
palavra:	<b>Strategic vision cannot be achieved without a coherent sustained industrial strategy</b>
origem:	Visão estratégica não pode ser alcançada sem estratégia industrial, apoiada coerente. A visão estratégica não pode ser conseguida sem uma estratégia industrial sustentada coerente. Não pode ser alcançada visão estratégica sem uma estratégia industrial contínua coerente.
de tradução:	2
palavra identificada no texto:	Adjetivos: Escolha Lexical Etião: Determinante - Adigo
<b>Frase Original:</b>	<b>The Treaty could mean a lot for the development of a truly united Single European Market.</b>
palavra:	inútil
origem:	BNC
PO3 Frase Original:	<b>AVB AD AD AD AD NMI</b>
palavra:	<b>The Treaty could mean a lot for the development of a truly united Single European Market</b>
origem:	O Tratado pode querer dizer muito para o desenvolvimento de um único Mercado Europeu verdadeiramente unido. O tratado pode significar muito para o desenvolvimento de um único mercado europeu verdadeiramente unido. O Tratado poderia significar muito para o desenvolvimento de um Único Mercado de europeu verdadeiramente unido.
de tradução:	3
palavra identificada no texto:	Ordem: Itador do DN



Frame Original:	Kasparov opted for maintaining a fluid yet stifling barrier of pawns in the centre of the board.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	AT9 A30 yet A30 N90
verso Original:	a fluid yet stifling barrier
translation:	O Kasparov optou para manter um fluido ainda barreira de stifling de peões no centro da tábua. *
en:	Kasparov opted mantendo um líquido contudo barreira stifling dos peões no centro da tábua. *
over:	Kasparov optou para manter uma barreira ainda sufocante fluida de peões no centro da tábua. *
pt:	Kasparov optou por ainda manter um fluido barreira sufocante de peões no centro da tábua. *
no de traduções:	4
como identificado na sêrie:	Adjectivos: Escolha Lexical Conjunções: Escolha Lexical Ordem: Interior do SN
reparar sobre a sêrie:	Kasparov optou por manter uma barreira fluida, mas/contudo cerrada de peões no centro do tabuleiro.
Frame Original:	We followed his hysterical yet provoking series on losses.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	POS A30 yet A30 N90
verso Original:	his hysterical yet provoking series
translation:	Seguimos suas série ainda provocando histórica em losses. *
en:	Não seguimos suas séries hysterical contudo provocando em losses. *
over:	Seguimos a sua série ainda provocante histórica losses. *
pt:	Não seguimos as séries ainda provocando históricas dele em losses. *
no de traduções:	4

Frame Original:	The Kent Care Project envisaged a key worker for particularly frail elderly people in the community.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	AV0 A30 A30 N90
verso Original:	particularly frail elderly people
translation:	O envisaged de Projeto de Cuidado de Kent um trabalhador chave para pessoas idosas particularmente fráguas na comunidade. *
en:	O projeto do cuidado de Kent envisaged um trabalhador chave para pessoas idosas particularmente fráguas na comunidade. *
over:	O projeto Kent Care contemplou um trabalhador chave para gente idosa particularmente dedicada na comunidade.
pt:	Translation Service not available.
no de traduções:	2
como identificado na sêrie:	Substantivos: Escolha Lexical Adjectivos: Plural dos Adjectivos
reparar sobre a sêrie:	Trad: O Projecto Kent Care vitima a disponibilização de um trabalhador para as pessoas idosas particularmente fráguas da comunidade.
Frame Original:	This model was initially based on the fairly crude statistical data available a century ago.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	AT9 AV0 A30 A30 N90
verso Original:	the fairly crude statistical data
translation:	Este modelo inicialmente foi baseado nos dados anatómicos claramente brutos disponível há muito e muito século. *
en:	Este modelo foi baseado inicialmente nos dados anatómicos razoavelmente crus disponíveis um século há. *
over:	Este modelo foi fundado inicialmente sobre os dados anatómicos bastante grosseiros disponíveis um século atrás.
pt:	Este modelo estava inicialmente um século atrás baseado nos dados anatómicos bastante

Frame Original:	The NFCC published an eleven volume historical series entitled Eras of Nonconformity.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	AT9 A30 A30 N90
verso Original:	an eleven volume historical series
translation:	O NFCC publicou um onze volume série histórica Eras intituladas de Nonconformity. *
en:	O NFCC publicou uma série histórica de onze volumes intitulada eras de nonconformity. *
over:	O NFCC publicou uma série de eleven-volume histórica eras intituladas de Nonconformity. *
pt:	O NFCC publicou um onze-volume que séries históricas intitularam Eras de Nonconformity. *
no de traduções:	4
como identificado na sêrie:	Adjectivos: Escolha Lexical Ordem: Interior do SN
reparar sobre a sêrie:	Trad: A NFCC publicou uma série histórica em onze volumes intitulada Eras de Não Conformidade.
Frame Original:	The sad reality for the newly elected Labour Minister is that Mr Patten's estimate was right.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	AT9 A30 A30 N90
verso Original:	the newly elected Labour Minister
translation:	A realidade triste para o Ministro recentemente eleito de Trabalho é aquela estimativa de Patten de Si era direita. *
en:	A realidade sad para o ministro labour novo eleito é que a estimativa do Sr. Patten era direita. *
over:	A realidade triste para o ministro de trabalho recentemente eleito é essa Senhor que a estimativa de Patten era constantemente. *
pt:	A realidade triste para os recentemente elegeram Labura o Ministro é a estimativa daquele Mr Patten era certa. *

Frame Original:	All military personnel at Lakenheath are tested, but high risk medical staff are given top priority.
ador:	indefinito
in:	BNC
do PO2 Frame Original:	A30 A30 N90
verso Original:	high risk medical staff
translation:	Todo pessoal militar em Lakenheath são testados, mas alto risco pessoal médico são dados prioridade de topo. *
en:	Tudo o pessoal militar em Lakenheath é testado, mas a equipe de funcionários médica high-risk é dada a prioridade superior. *
over:	Todos as pessoas militares a Lakenheath são testados, mas as high-risk pessoal médico é dada a prioridade superior. *
pt:	Translation Service not available.
no de traduções:	3
como identificado na sêrie:	Substantivos: Escolha Lexical Adjectivos: Escolha Lexical Ordem: Interior do SN
reparar sobre a sêrie:	Trad: Todo o pessoal militar em Lakenheath é testado, mas a prioridade máxima é dada ao pessoal médico de elevado risco.
Frame Original:	This avoids the long unpredictable legal proceedings that follow an attempt to lay off workers.
ador:	indefinito
in:	PERIODICA
do PO2 Frame Original:	AT9 A30 A30 A30 N90
verso Original:	the long unpredictable legal proceedings
translation:	Isto evita os processos legais, imprevisíveis longos que seguem uma tentativa despedir trabalhadores. *
en:	Isto evita as contuações legais imprevisíveis longas que seguem uma tentativa de se encontrar fora trabalhadores. *
over:	Acceso: Faltado
pt:	Isto evita os socorridos legais imprevisíveis longos que seguem uma tentativa para demiti

Lexical Bundles

EXAMPLES

- Now let us look at some examples:
  - Homographs
  - Polysemy
  - Complex noun phrases
  - Lexical bundles

<b>Case Original:</b>	<b>I want you to be concerned about your next door neighbors.</b>
Por:	incluindo
Por:	ManuWEB
P201 Phrase Original:	<b>I want you to VBI AB PFP</b>
Mo Oribio:	<b>I want you to</b>
Tradução:	Quero ser preocupado sobre seu ou lado vizinho. *
er:	Eu quero-o ser preocupado sobre seu vizinho seguinte da porta. *
er:	Acesso Falhado
er:	Eu quero que você seja interessado sobre seu próximo vizinho de porta.
de Tradução:	2
er Identificador no de:	Vérbo: Escolha Lexical
er:	Conjunções: Conjunções Subordinadas
er:	Modo verbal: Infinitivo no lugar de conjunção
<b>Case Original:</b>	<b>What do you want me to do?</b>
Por:	incluindo
Por:	COMPARA
P201 Phrase Original:	<b>OTO VMO you want me to VDI UNC</b>
Mo Oribio:	<b>you want me to</b>
Tradução:	O que quer me fazer? *
er:	Que você me quer fazer? *
er:	Acesso Falhado
er:	O que quer você eu fazer? *
de Tradução:	3
er Identificador no de:	Pronomes: Pronomes Pessoais (Função Sujeito)
er:	Conjunções: Conjunções Subordinadas
er:	Modo verbal: Infinitivo no lugar de conjunção
er:	Pronomes: Pronomes (objetos, "object", "object")

<b>Case Original:</b>	<b>I'm not sure if I'm meant to be a factory worker.</b>
Por:	incluindo
Por:	?
P201 Phrase Original:	<b>CJS I'm meant to VBI ATB NNI</b>
Mo Oribio:	<b>I'm meant to</b>
Tradução:	Eu não estou seguro se eu quando dizer para ser trabalhador de fábrica. *
er:	Eu não sou certo se eu for significado ser um trabalhador de fábrica. *
er:	Não tenho certeza se eu quando dizer para ser um trabalhador de fábrica. *
er:	Eu não estou seguro se eu quando dizer ser um trabalhador de fábrica. *
de Tradução:	4
er Identificador no de:	Vérbo: Escolha Lexical
er:	Novamente a preposição "Para" em excesso, desta vez em duas versões: FreeTranslation e E-T Sense.
er:	A conjunção "Se" obriga novamente a recorrer ao contexto, para determinar a necessidade ou não do uso do conjunção.
<b>Case Original:</b>	<b>You're meant to do that stuff in private.</b>
Por:	incluindo
Por:	ManuWEB
P201 Phrase Original:	<b>You're meant to VDI DTB NNI PFP</b>
Mo Oribio:	<b>You're meant to</b>
Tradução:	São quando dizer para fazer aquele material em privado.
er:	Você é significado fazer esse material em confidencial. *
er:	Você é significado para fazer esse material em particular. *
er:	Você é quando dizer fazer aquela matéria-prima dentro privado. *
de Tradução:	3
er Identificador no de:	Vérbo: Escolha Lexical
er:	Preposição "Para" em excesso

<b>Case Original:</b>	<b>In the course of the trial one Glazkov made a false confession which he later retracted.</b>
Por:	Jilma
Por:	UNC
P201 Phrase Original:	<b>PSP ATB NNI PFP</b>
Mo Oribio:	<b>in the course of</b>
Tradução:	No curso do julgamento um Glazkov fez uma confissão falsa que ele mais tarde retraiu. *
er:	No curso do um Glazkov experimental fez uma confissão falsa que retraiu mais tarde. *
er:	No curso da prova um Glazkov fez uma confissão falsa que ele depois retraiu. *
er:	No curso do um Glazkov de ensaio fez uma falsa confissão que ele retraiu depois. *
de Tradução:	4
er Identificador no de:	Substantivos: Escolha Lexical
er:	Durante o julgamento, Glazkov fez uma confissão falsa de que depois se retractou.
er:	Análise de associações lexicas
er:	"Course" é traduzido à letra, ignorando o significado global da expressão lexicalizada
<b>Case Original:</b>	<b>By the time of the divorce in 1996, the assets of the family were considerable.</b>
Por:	Jilma
Por:	UNC
P201 Phrase Original:	<b>ATB NNI PFP ATB</b>
Mo Oribio:	<b>the time of the</b>
Tradução:	Pelo tempo do divórcio em 1996, os ativos da família eram consideráveis. *
er:	Pela época do divórcio em 1996, os recursos da família eram consideráveis. *
er:	Pelo tempo do divórcio em 1996, o ativo da família era considerável. *
er:	Até que do divórcio em 1996, os ativos da família fossem consideráveis. *
de Tradução:	4
er Identificador no de:	Substantivos: Escolha Lexical
er:	Na altura do divórcio, em 1996, os bens da família eram consideráveis.

<b>Case Original:</b>	<b>The rate of change in the private sector is incredibly great.</b>
Por:	Jilma
Por:	UNC
P201 Phrase Original:	<b>ATB NNI PFP NNI PFP</b>
Mo Oribio:	<b>the rate of change of</b>
Tradução:	O índice de mudança no setor privado é incrivelmente grande. *
er:	A taxa de mudança no setor confidencial é incrivelmente grande. *
er:	A taxa de mudança ao setor privado é incrivelmente grande. *
er:	A taxa de mudança no setor privado é incrivelmente grande. *
de Tradução:	4
er Identificador no de:	Substantivos: Escolha Lexical
er:	A rapidez das alterações no sector privado é extraordinariamente elevada.
er:	Análise de associações lexicas
er:	A expressão "Rate of change" pode muito bem ser encarada como um termo, com várias traduções diferentes de acordo com o área de conhecimento em causa (velocidade de variação, taxa de variação). Nos sistemas de tradução automática, o termo não é reconhecido e mesmo a escolha lexical da primeira palavra é claramente errada em qualquer dos casos.
<b>Case Original:</b>	<b>The aim of this meeting was to go one stage further.</b>
Por:	Jilma
Por:	UNC
P201 Phrase Original:	<b>ATB NNI PFP NNI NNI</b>
Mo Oribio:	<b>the aim of this meeting</b>
Tradução:	O objetivo desta reunião era ir mais uma etapa mais adiante. *
er:	O alvo desta reunião era ir mais mais um estágio.
er:	O objetivo desta reunião era ir uma fase ulteriore. *
er:	A finalidade desta reunião era ir um estágio mais adiante.
de Tradução:	2
er Identificador no de:	Substantivos: Escolha Lexical

<b>Case Original:</b>	<b>The Council must of course proceed on the basis of the veto.</b>
Por:	Jilma
Por:	?
P201 Phrase Original:	<b>PFP ATB NNI PFP</b>
Mo Oribio:	<b>on the basis of</b>
Tradução:	O Conselho deve naturalmente proceder pelo voto. *
er:	O conselho deve naturalmente prosseguir na base do voto. *
er:	O conselho deve proceder naturalmente segundo o voto.
er:	O Conselho tem que proceder claro que em base do voto. *
de Tradução:	3
er Identificador no de:	Preposições: Escolha Lexical
er:	O conselho deve naturalmente proceder com base no voto
er:	Análise de associações lexicas
er:	Trata-se de uma estrutura muito comum, sendo que a preposição em Português é diferente, algo que os sistemas de tradução não identificaram.
<b>Case Original:</b>	<b>In 1926 38.3 percent of the population were Ukrainians against 37.9 at the time of the 1923 census.</b>
Por:	Jilma
Por:	UNC
P201 Phrase Original:	<b>PFP ATB NNI PFP</b>
Mo Oribio:	<b>at the time of</b>
Tradução:	Em 1926 38.3 por cento da população eram ucranianos contra 37.9 no tempo do 1923 censo. *
er:	Em 1926 38.3 por cento da população eram ucranian contra 37.9 no momento do 1923 censo. *
er:	Em 1926 38.3 % da população estava o ucraniano contra 37.9 no momento do 1923 censo. *
er:	Em 1926 38.3 por cento da população eram ucranianos contra 37.9 na hora do 1923 censo. *
de Tradução:	3
er Identificador no de:	Substantivos: Escolha Lexical

<b>Case Original:</b>	<b>They should be given a brief personal experience of the nature of prison life.</b>
Por:	Jilma
Por:	?
P201 Phrase Original:	<b>PFP ATB NNI PFP</b>
Mo Oribio:	<b>of the nature of</b>
Tradução:	Devem ser dados uma experiência pessoal breve da natureza de vida da prisão. *
er:	Devem ser dados uma experiência pessoal breve da natureza da vida da prisão.
er:	Devem-ter-se ser dada uma experiência pessoal breve da natureza de vida de prisão. *
er:	Devem ser dados uma experiência pessoal breve da natureza de vida de prisão. *
de Tradução:	3
er Identificador no de:	Preposições: Contracção com Antigos
er:	Devem ser dadas uma breve experiência pessoal das características da vida na prisão.
er:	Análise de associações lexicas
er:	A preposição que aparece a seguir a "natureza" deve ser corrigida com o artigo definido. Por outro lado, é discutível a escolha do termo "natureza" em termos estilísticos.
<b>Case Original:</b>	<b>This value seems to be disproportionately in view of the simplified regime promised.</b>
Por:	Jilma
Por:	ManuWEB
P201 Phrase Original:	<b>PFP NNI PFP ATB</b>
Mo Oribio:	<b>in view of the</b>
Tradução:	Este valor parece a nós desproporcional em vista do regime simplificado prometido. *
er:	Este valor parece desproporcional na vista do regime simplificado prometido. *
er:	Este valor não parece desproporcional em vista do regime simplificado prometido. *
er:	Este valor parece a nós desproporcional devido ao regime simplificado prometido.
de Tradução:	3
er Identificador no de:	Preposições: Escolha Lexical
er:	Este valor parece-nos desproporcional, tendo em contãem vista o regime simplificado que foi prometido.

Forward	
<b>base Original:</b>	<b>He appealed for leniency on the grounds that nothing had been touched.</b>
cor:	<b>Jimia</b>
	<b>UNC</b>
PDS Phrase Original:	<b>PSP ATB INO CST</b>
no Châno:	<b>on the grounds that</b>
translation:	Apelo para indulgência nos chãos esse nada tem sido tocado. *
	Apelo para o lenency nas terras em que nada tinha sido tocado. *
cor:	Ele pediu lenency nas terras que nada tinha sido compido. *
	Ele atriu para indulgência nos chãos que nada tinha sido tocado. *
de tradução:	4
nos identificamos na:	Substantivos: Escolha Lexical
do:	
do proposto:	Ele apelo à indulgência alegando que ninguém tinha tocado em nada.
do termo extra frase de:	Análise de associações léxicas
do termo gerado:	Os motores de tradução não reconhecem a estrutura e a tradução é sempre errada. a expressão como um todo deve ser traduzida por "alegado", "com base em", etc.
<b>base Original:</b>	<b>To the extent that X is similar to Y, the acquired equivalence will operate.</b>
cor:	<b>Jimia</b>
	<b>UNC</b>
PDS Phrase Original:	<b>PSP ATB INO CST</b>
no Châno:	<b>to the extent that</b>
translation:	A extensão que X é semelhante a Y, a equivalência adquirida operará. *
	Até ao ponto em que X é similar a Y, a equivalência adquirida operará. *
cor:	A extensão que X é similar a Y, a equivalência comprada operará. *
	Para a extensão que X é semelhante a Y, a equivalência adquirida operará. *
de tradução:	4
nos identificamos na:	Substantivos: Escolha Lexical
do:	Substantivos: Escolha Lexical