

# CONTROLO DA QUALIDADE LINGUÍSTICA E PARÁFRASES

-

## AJUDANTES DA TRADUÇÃO AUTOMÁTICA

Anabela Barreiro

LINGUATECA/FCCN

CLUP

**Aspectos do Processamento da Linguagem Natural em português**

**Ferramentas e recursos existentes > investigação futura**

# CONTEÚDO DA SESSÃO

- ❑ Controlo da qualidade linguística
  - Ferramentas
- ❑ Paráfrase
  - Definição
  - Funções
  - Bases teóricas
  - Tipos
  - Aplicações (importância em PLN e outras áreas)
- ❑ Paráfrases de construções com verbo suporte
  - Sistemas e bases de conhecimento parafrástico adquirido a partir do léxico

# CONTEÚDO DA SESSÃO

- ❑ Paráfrases e tradução
  - Tradução de construções com verbo suporte
  - Uso e impacto de paráfrases em TA
    - Tese
    - Objectivos
    - Avaliação
    - Conclusão
    - Aplicação: ParaMT – parafraseador para a TA
    - TA com paráfrases
  
- ❑ Referências

# CONTROLO DA QUALIDADE LINGUÍSTICA

- Técnicas de verificação da qualidade da documentação especializada
- Verificação ortográfica, gramatical, estilística e terminológica dos documentos técnicos
- Linguagem consistente, directa e simples
- Gramática restrita (evita certos tipos de construções)
- Evita raciocínios complexos, figuras de estilo, metáforas, etc.
- Optimização de texto
- Métodos de parafraseamento
- Redução de palavras

# CONTROLO DA QUALIDADE LINGUÍSTICA

- ❑ Tem sido utilizada essencialmente na escrita técnica – em manuais, guias de utilização, textos para a tradução automática, entre outros.
  - Manuais com instruções em como operar maquinaria, automóveis
  - Guias de utilização de aparelhos eléctricos, programas informáticos
- ❑ As técnicas de controlo de qualidade linguística são extensíveis a:
  - Relatórios médicos
  - Sistemas de detecção e denúncia de eventos adversos ou farmacovigilância
  - Artigos científicos
  - Gestão de conteúdos para empresas
  - Gestão de conteúdos dos sítios na internet
  - Sistema de emissão de mensagens e aplicações de diálogo imediato (SMS, messenger, e outros serviços de tecnologia de comunicação)
  - E mais...

# FERRAMENTAS E LINGUAGEM CONTROLADA

**MULTIDOC** - [Haller, 2000]

**KANT CE Checker** – [Mitamura and Nyberg, 2001]; [Mitamura et al., 2003]

**CLAT** – [Schmidt-Wigger, 1998]; [Haller, 2000]; [Reuther and Schmidt-Wigger, 2000]

**CLOUT** - the Controlled Language Optimized for Machine Translation. [Uwe Muegge, 2006] <http://www.muegge.cc/controlled-language.htm>

## Linguagem controlada

Agilent Technologies English (Smartny, 2002), Attempto Controlled English (ACE) (Fuchs and Schwitter, 1996), Boeing Technical English (Wojcik, Holmback, and Hoard, 1998), Caterpillar Technical English (Kamprath, Adolphson, Mitamura, and Nyberg, 1998), Diebold Controlled English (Moore, 2000), Ericsson English (Ericsson, 2000), General Motors CASL (Means and Godden, 1996), Global English (Means, Chapman and Liu, 2000), Kodak English (Kodak, 2000), Nortel Standard English (Smartny, 2002), Océ Technologies English (Smartny, 2002), Perkins PACE (Douglas and Hurst, 1996) and Xerox MCE (Xerox, 2001).

# PARÁFRASES

- ❑ Riqueza linguística = muitas formas de “dizer a mesma coisa” usando palavras, expressões ou frases diferentes
- ❑ Seres humanos - usam estratégias linguísticas diferentes para expressar conhecimento do mundo ou de situações do mundo real (**paráfrases referenciais** - relacionadas com factores sócio-culturais)
- ❑ Expressões alternativas (sintaxe ≠) para expressar o mesmo significado (semântica =)
- ❑ Retêm “equivalência conceptual aproximada” [Barzilay, 2003]
- ❑ Parafraseamento - técnica para articular uma ideia ou veicular informação de forma alternativa, facilitando a sua compreensão e e para tornar os textos mais valiosos para a sua audiência.

# FUNÇÕES

- ❑ **Explicar** - definir determinado termo ou conceito inexistente num país ou realidade sócio-cultural diferente – referência
- ❑ **Interpretar** - parafrasear termos técnicos ou conceitos/ideias complexas em palavras ou expressões mais simples de entender para um falante comum, não especialista e menos familiarizado com certas terminologias ou complexidade linguística
- ❑ **Clarificar** - comunicar de forma mais clara – melhorar a compreensibilidade
- ❑ **Simplificar** - ideias e conceitos usando uma gramática mais fácil, eliminando repetições e o que não acrescenta significado novo
- ❑ **Encurtar** - tornar um texto mais sucinto, conciso e menos palavroso
- ❑ **Valorizar** - enriquecer a frase ou texto sem o tornar mais complexo
- ❑ **Mudar o estilo** - estabelecer um estilo diferente, mais ou menos formal, mais ou menos directo, metafórico, consistente, uniforme, etc.
- ❑ **Praticar a escrita** - usado como exercício pedagógico
- ❑ **Traduzir** – transpor para outra língua de forma a manter a fluência nessa língua

# BASES LINGUÍSTICAS

- **Gramática Generativa Transformacional** [Chomsky, 1957] - as cadeias de expressões linguísticas são consideradas estruturas superficiais que podem formar-se com a ajuda de um conjunto de regras de formação a partir de estruturas subjacentes. A cada estrutura profunda pode fazer-se corresponder depois uma classe de estruturas superficiais que são como que paráfrases dela (ex: transformação activa-passiva).
- **Teoria Sentido-Texto** [Mel'čuk, 1988, 1996, 2003] – as construções com verbos suporte são descritas e formalizadas in termos de funções lexicais (ex: Magn(X) faz o mapeamento de uma palavra X em palavras que a intensifica - Magn("shave")='clean'). As funções lexicais podem ser usadas como uma interlíngua para facilitar a transferência lexical em projectos de TA. [Milićević, 2007] estabelece uma tipologia de paráfrases baseadas em regras léxico-sintácticas.
- **FrameNet** [Fillmore et al., 2002, 2003] – regista a informação necessária para a representação de relações entre argumentos, nomeadamente entre um **verbo suporte** e uma nominalização. Explanatory Combinatorial Dictionary
- **NomBank** [Meyers et al., 2004b, 2004b] – mapeia posições sintácticas (argumentos do predicado e nominalizações) e identifica os complementos de cada nominalização, relacionando os complementos nominais aos argumentos do verbo correspondente, incluindo informação acerca de **verbos suporte**.
- **Teoria do Léxico-Gramática** [M. Gross, 1975, 1981, 1996 e seguidores] – quadro teórico estabelecido na gramática transformacional de Harris. A unidade básica de análise e descrição do significado é a frase simples ou elementar (o predicado ou núcleo da frase e os seus argumentos obrigatórios). Os predicados podem ser nomes, adjectivos ou advérbios. Por exemplo, os nomes predicativos seleccionam os seus argumentos e os verbos suporte. As transformações são relações não orientadas de equivalência entre frases (paráfrases).

# TIPOS DE PARÁFRASE

❑ Lexical (sinónimo) versus sintáctica

❑ Granularidade

- **Palavra**

*apresentar/expôr (um trabalho)*

*O João comprou um carro/automóvel*

- **Expressão**

*uma pintura a óleo > uma pintura feita a óleo*

*dar apoio a > apoiar*

- **Frase**

*O João deu um presente à Maria > a Maria recebeu um presente do João*

*O João abraçou a Maria > a Maria foi abraçada pelo João*

[adapt. Rodriguez, 2009]

# TIPOS DE PARÁFRASE

□ Exemplos de [Barzilay, 2003] para o inglês

- **Palavra**

*debate* ≡ *discuss*

- **Expressão**

*wooden frame* ≡ *frame made of wood*

*to aim the guns* ≡ *to get the best firing angles.*

- **Frase**

*Emma did not know how to waltz* ≡ *Emma had no clue about waltzing*

*The paper was hotly debated, causing a fine old uproar*

≡ *The article was warmly discussed, which procured it a high reputation.*

*Eli planted a tomato bush.* ≡ *a tomato bush was planted by Eli.*

*Louis sold the book to Noemie.* ≡ *Noemie bought the book from Louis.*

[adapt. Rodriguez, 2009]

# TIPOS DE PARÁFRASE

- ❑ Paráfrases idiossincráticas versus paráfrases produtivas
  - Atômicas versus composicionais (regras composicionais representadas como árvores de dependência parcialmente lexicalizadas – ex: activa-passiva - SN1 VB1 SN2 > SN2 foi VB1-ado por SN1)
- ❑ Efeitos de distorção do significado [Dras, 1999]
  - Mudança de perspectiva
  - Mudança de ênfase
  - Mudança de relação
  - Apagamento
  - Movimento (deslocação) de estruturas
- ❑ Classificações de dimensões múltiplas
  - [Dras, 1999]
  - [Fujita, 2005]
  - [Boonthum, 2002]

[adapt. Rodriguez, 2009]

# APLICAÇÕES

- ❖ **Resposta a pergunta** – importante na validação de respostas - indicam que várias respostas equivalentes (mas não iguais) estão correctas - [Ibrahim et al., 2003], [Paşca, 2003], [Duboué & Chu-Carroll, 2006]
- ❖ **Extracção de informação e prospecção de texto** – ajudam nas tarefas de categorização de texto ou de mapeamento de textos com características semelhantes - [Ibrahim et al., 2003], [Shinyama et al., 2002] [Shinyama & Sekine, 2003], [Sekine, 2005] [Paşca, 2005], [Paşca & Dienes, 2005]
- ❖ **Sumarização** – permitem que informação em vários documentos seja condensada e ajudam a melhorar a qualidade dos resumos - [McKeown et al., 2002], [Barzilay, 2001, 2003], [Hirao et al., 2004] [Zhou et al., 2006b]
- ❖ **Geração de linguagem natural** – aumentam o poder expressivo dos sistemas de geração ao permitirem a produção de texto mais variado e fluente - [Iordanskaja et al. 1991]
- ❖ **Tradução (automática)** – ajudam a criar traduções mais fluentes, mais claras e fáceis de entender. São um elemento importante na avaliação dos resultados da TA, ao permitirem comparar traduções equivalentes - [Zhou et al., 2006], [Callison-Burch et al., 2006a, 2006b, 2007 and 2008].

# APLICAÇÕES

- ❖ **Controlo da qualidade linguística** – permitem desambiguar, eliminar redundâncias, simplificar e uniformizar a produção de frases e criar um texto de maior compreensibilidade
- ❖ **Produção e pré-edição de texto** - encurtam texto (redução do número de palavras) para a TA e oferecem alternativas que melhoram o significado das frases do texto da língua de partida
- ❖ **Redacção estilística** – eliminam redundâncias ou escolhas estilísticas inapropriadas, apresentando outros substitutos mais adequados e objectivos
- ❖ **Detecção de plágio** – permitem detectar frases copiadas e ligeiramente transformadas
- ❖ **Ensino da língua** – ajudam no ensino e aprendizagem da língua materna e de línguas estrangeiras

# PARÁFRASES DE CONSTRUÇÕES COM VERBO SUPORTE

**CONSTRUÇÃO COM VERBO SUPORTE** = unidade lexical multpalavra ou predicado complexo constituído por um verbo semanticamente fraco (o verbo suporte) e um nome, adjectivo ou advérbio predicativo (o predicado da frase).

Os nomes predicativos podem ser:

- **Morfologicamente derivados de um verbo**

PT: fazer uma apresentação de N = apresentar N

- **autónomos**

PT: fazer um mestrado - \*mestrar

# PARÁFRASES DE CONSTRUÇÕES COM VERBO SUPORTE

## FORMALIZAÇÃO AO NÍVEL DO DICIONÁRIO

- ❑ Marcação explícita da derivação, associações semânticas entre predicados (verbos, nomes predicativos derivados e autónomos, adjectivos predicativos) e ligações entre predicados e verbos suporte por eles seleccionados:

ex: **dar um abraço a X – abraçar X**

**visitar - fazer uma visita – efectuar uma visita – realizar uma visita**

**cansar-se – ficar cansado**

**estudar X - tirar um curso de X**

- ❑ Associações semânticas entre adjectivos e advérbios

ex: **ter uma evolução gradual – evoluir gradualmente**

# PARÁFRASES DE CONSTRUÇÕES COM VERBO SUPORTE

## FORMALIZAÇÃO AO NÍVEL DO DICIONÁRIO

Associações semânticas entre verbos e construções com verbo suporte com *passar*

**revistar**,V+FLX=FALAR+Aux=1+OBJTRundif75+Subset=319+EN=**search**+VSUP=passar+DRV=NDRV16:CASA+PrepN=a

**escovar**,V+FLX=FALAR+Aux=1+OBJTR19+Subset=654+EN=**brush**+VSUP=passar+DRV=NDRV332:CASA+PrepN=em +PrepN=por

**pentear**,V+FLX=PASSEAR+Aux=1+OBJTRundif92+Subset=184+EN=**comb**+VSUP=passar+DRV=NDRV73:ANO +PrepN=em

**engraxar**,V+FLX=FALAR+Aux=1+INOPwith+Subset=237+EN=**polish**+VSUP=dar+DRV=NDRV553:CASA+PrepN=a +VSUP=passar+PrepN=em

**pintar**,V+FLX=FALAR+Aux=1+INOPwith+Subset=201+EN=**paint**+VSUP=fazer+DRV=NDRV109:ANO+VSUP=passar+DRV=NDRV571:CASA+PrepN=em

**engomar**,V+FLX=FALAR+Aux=1+INOPmisc29+Subset=125+EN=**iron**+VSUP=passar+Prep=a+DRV=NDRV572:ANO +PrepN=em+PrepN=sobre

# PARÁFRASES DE CONSTRUÇÕES COM VERBO SUPORTE

que Edward pudesse	dar início à/iniciar	carreira de clínico geral
que Edward pudesse	dar início à/começar	carreira de clínico geral
te, se me é permitido	dar uma opinião/ter uma opinião	. Bernard aproveitou p
te, se me é permitido	dar uma opinião/dar uma opinião	. Bernard aproveitou p
te, se me é permitido	dar uma opinião/opinar	. Bernard aproveitou p
h. -- Quanto se deve	dar de gorjeta/gratificar	? -- O que quis dizer f
o que, se puder, devo	dar uma ajuda ao/ajudar	Council. Um deles abo
isso mandaram-no ir	dar uma volta/passear	. Uma mulher com a sa
de Persse, começa a	dar exemplos/exemplificar	: O Espelho da Sincer
racional impedia-a de	dar voz ao/ouvir	medo. Ainda não eram
racional impedia-a de	dar voz ao/escutar	medo. Ainda não eram
e sentia-se capaz de	dar um murro em/esmurrar	quem tentasse detê-lo
l. Gostávamos de lhe	dar uma palavrinha/falar	. -- Então vamos dar u
rinha. -- Então vam...	dar um aperto de/apertar	mão. «Deixar que o se
órios, «poderia não...	dar lugar a uma/originar	colisão qualquer, mas
órios, «poderia não...	dar lugar a uma/causar	colisão qualquer, mas
to? Às vezes ouvia-a	dar instruções a/instruir	Julie acerca das comp
a do altruísmo, podia	dar sempre origem a/originar	confusões. O Oliver g
um tractor. -- Posso	dar uma olhadela/olhar	? -- perguntou Zoe. E
ai que gostava de lhe	dar uma palavrinha/falar	, está bem? Harriet lev
ando eu não estiver a	dar aulas/leccionar	, e nos fins-de-semana
ando eu não estiver a	dar aulas/ensinar	, e nos fins-de-semana
u com o charuto para	dar ênfase às/enfatizar	suas palavras. «Gosta
Agatha e Félix foram	dar um passeio pelo/passear	campo e deixaram o v
Agatha e Félix foram	dar um passeio pelo/andar	campo e deixaram o v
e fui incapaz de lhe	dar resposta/responder	. Apenas me preocupa
u-lhe: Não precisa de	dar explicações/explicar	. Reza fita-me como se
se me permite, posso	dar uma sugestão/fazer uma sugestão	.
se me permite, posso	dar uma sugestão/dar uma sugestão	.
se me permite, posso	dar uma sugestão/sugerir	.

# SISTEMAS E BASES DE CONHECIMENTO PARAFRÁSTICO ADQUIRIDO A PARTIR DO LÉXICO

**Para o português:**

**ReEscreve** – [Barreiro , 2008] - <http://www.linguateca.pt/ReEscreve/>

**Para outras línguas:**

**DIPRE** - Dual Iterative Pattern Relation Expansion – [Brin, 1999]

<http://www.alexmayers.com/projects/DIPRE/>

**Snowball** [Agischtein, Gravano, 2001]

**DIRT** (Discovery of Inference Rules from Text) – [Lin and Pantel , 2001]

<http://demo.patrickpantel.com/Content/LexSem/paraphrase.htm>

[Ravichandran, Hovy, 2002]

[Bhagat, Ravichandran, 2008]

# PARÁFRASES E TRADUÇÃO

- ❑ As paráfrases são muito importantes numa língua, mas são ainda mais importantes na tradução de uma língua para outra.
- ❑ Entender o texto antes de traduzir implica muitas vezes parafraseá-lo de modo a que seja mais facilmente entendido na língua de destino.
- ❑ Línguas diferentes usam formas de expressão diferentes e apresentam um comportamento sintáctico-semântico diferente.
- ❑ Palavras e expressões podem ser ambíguas e ter mais do que um sinónimo e, como tal, mais do que uma tradução possível.
- ❑ A tradução implica sinonímia entre línguas diferentes (ex. period > período = época, quando diz respeito a uma realidade temporal)
- ❑ A tradução literal nem sempre é possível ou desejada. Por vezes, não existem equivalentes directos entre línguas.
- ❑ Nem sempre é possível encontrar um sinónimo na língua de destino, porque existem diferentes realidades conceptuais e referenciais. Nesses casos, a tradução precisa de ser mais dinâmica e menos “agarrada” ao original.

# TRADUÇÃO DE CONSTRUÇÕES COM VERBOS SUPORTE

fazer adiamentos => to give extensions

fazer perguntas => to ask questions

fazer o exame => to take the exam

fazer uma visita => to pay a visit

fazer uma operação/cirurgia => to have an operation/surgery

Construções com verbo suporte em que o verbo suporte não pode ser traduzido literalmente

fazer publicidade => to advertise

fazer um esboço => to draft

fazer compras => to shop

fazer a barba => to shave

estar presente/ausente = attend/not attend

Construções com verbo suporte traduzidas naturalmente por verbos

# IMPACTO DAS PARÁFRASES NA TRADUÇÃO AUTOMÁTICA

## METRA3

Tradução de "Eles querem dar uma vista de olhos ao museu"

E qual a melhor tradução na sua opinião ?

- They are going to give a sight of eyes to the museum FreeTranslation
- They want to give an eye sight to the museum E-Translation Server
- They want to give a sight of eyes to the museum WorldLingo
- Nenhuma destas traduções

## METRA3

Tradução de "Eles querem ver o museu"

E qual a melhor tradução na sua opinião ?

- They want to see the museum WorldLingo
- They are going to see the museum FreeTranslation
- They want to see the museum E-Translation Server
- Nenhuma destas traduções

**SISTEMAS DE TA  
APRESENTAM TRADUÇÃO  
LITERAL DE UNIDADES  
LEXICAIS MULTIPALAVRA  
(CONSTRUÇÕES COM  
VERBOS SUPORTE)**

**AH... ASSIM ESTÁ  
MELHOR!**

**A SUBSTITUIÇÃO DE UMA EXPRESSÃO POR UM  
VERBO RESULTA NA SUPERIOR QUALIDADE DA  
TRADUÇÃO PARA DIFERENTES SISTEMAS**

# IMPACTO DAS PARÁFRASES E CONTROLO DE QUALIDADE LINGUÍSTICA NA TA DE FRASES COM CVS DA ÁREA BIOMÉDICA

<i>Por falta de condições técnicas, ele foi <u>removido</u> para o Hospital das Clínicas, onde se fez uma amputação a nível de ombro.</i>	
FreeTranslation	For lack of technical conditions, he was <u>removed</u> for the Hospital of the Clinics, where <u>was done an amputation</u> in terms of shoulder.
WorldLingo	Due to conditions techniques, it it was <u>removed</u> for the Hospital of the Clinics, where if the shoulder level <u>made an amputation</u>
<i>Por falta de condições técnicas, ele foi <u>transportado</u> para o Hospital das Clínicas, onde os médicos amputaram o seu braço ao nível do ombro.</i>	
FreeTranslation	Due to conditions techniques, it it was <u>carried</u> to the Hospital of the Clinics, where <u>the doctors had amputated its arm</u> to the level of the shoulder.
WorldLingo	For lack of technical conditions, he was <u>transported</u> for the Hospital of the Clinics, where <u>the doctors amputated his arm</u> level with the shoulder.
<i>Por falta de condições técnicas, ele foi <u>transportado</u> para o Hospital das Clínicas, onde o seu braço foi amputado ao nível do ombro.</i>	
FreeTranslation	for lack of technical conditions, he was <u>transported</u> for the Hospital of the Clinics, where <u>arm was amputated</u> level with the shoulder.
WorldLingo	Due to conditions techniques, it it was <u>carried</u> to the Hospital of the Clinics, where its <u>arm was amputated</u> to the level of the shoulder.

# TESE

O conhecimento linguístico acerca de paráfrases ajuda a melhorar a qualidade da TA

➤ A conversão de construções com verbos suporte em verbos ou expressões verbais morfossintáctica e/ou semanticamente relacionados e a especificação adequada e inclusão de informação acerca da estrutura argumental do predicado produz uma linguagem de melhor qualidade tanto no domínio geral como no domínio técnico e torna a TA mais fiável.

1. Reduz a complexidade (= parafraseamento por simplificação)
2. Reduz a ambiguidade lexical (a paráfrase veícula a interpretação correcta)
3. Facilita a interpretação quando a construção com verbo suporte é idiomática
4. Reduz o número de palavras (muitas vezes, melhorando a qualidade do texto a ser traduzido)
5. Melhora a qualidade da tradução

# OBJECTIVOS

- Criar uma base de conhecimento lexical, sintáctico e semântico à volta das construções com verbos suporte
- Aplicar este conhecimento linguístico ao parafraseamento – parafrasear e reescrever construções com verbos suporte do domínio geral e específico, tais como *fazer planos* ou *fazer uma operação* com equivalentes correctos. Os equivalentes podem ser:
  - Verbos, tais como *planificar* ou *operar*
  - Variantes léxico-sintácticas e estilísticas das construções com verbos suporte originais, tais como *realizar / efectuar planos* ou *realizar uma operação* ou *submeter-se a uma operação/cirurgia* ou *ser operado*), dependendo da estrutura argumental e do preenchimento lexico-semântico dos argumentos.
- Mostrar como estas paráfrases melhoram a qualidade do texto de origem
- Mostrar que os resultados da TA podem melhorar consideravelmente quando se utilizam capacidades parafrásticas

# AVALIAÇÃO

Impacto das paráfrases  
de construções com  
verbos suporte na  
tradução

5 verbos suporte

100 frases com construções com verbos suporte (para cada língua PT e EN)

100 frases com paráfrases correspondentes (para cada língua)

20 frases para cada verbo suporte – parafraseamento manual para cada uma destas frases  
> 20 frases com construções com verbos suporte + 20 paráfrases das frases originais

TOTAL: 200 frases para cada língua

100 com construções com verbos suporte + 100 com paráfrases

10 avaliadores traduziram e avaliaram as traduções

alunos do Mestrado em Tradução e Serviços Linguísticos

Resultado de qualidade equiparada	A tradução da CVS é melhor	A tradução da paráfrase é melhor
17%	26%	57%
19%	30%	51%

Tanto para EN-PT como para PT-EN, os avaliadores consideraram que as traduções das paráfrases apresentavam uma melhor qualidade.

Os resultados da avaliação confirmam que este tipo de paráfrases ajuda os sistemas de TA a produzir melhores resultados.

# CONCLUSÃO

**O conhecimento linguístico formalizado neste estudo quando aplicado à TA ajuda a melhorar a qualidade da tradução.**

Quando uma CVS é identificada e substituída com expressões verbais semanticamente equivalentes como tarefa de pré-processamento:

verifica-se uma **melhoria de 21%** na qualidade dos resultados avaliados de tradução automática PT-EN

e uma **melhoria de 31%** na qualidade dos resultados avaliados de tradução automática EN-PT.

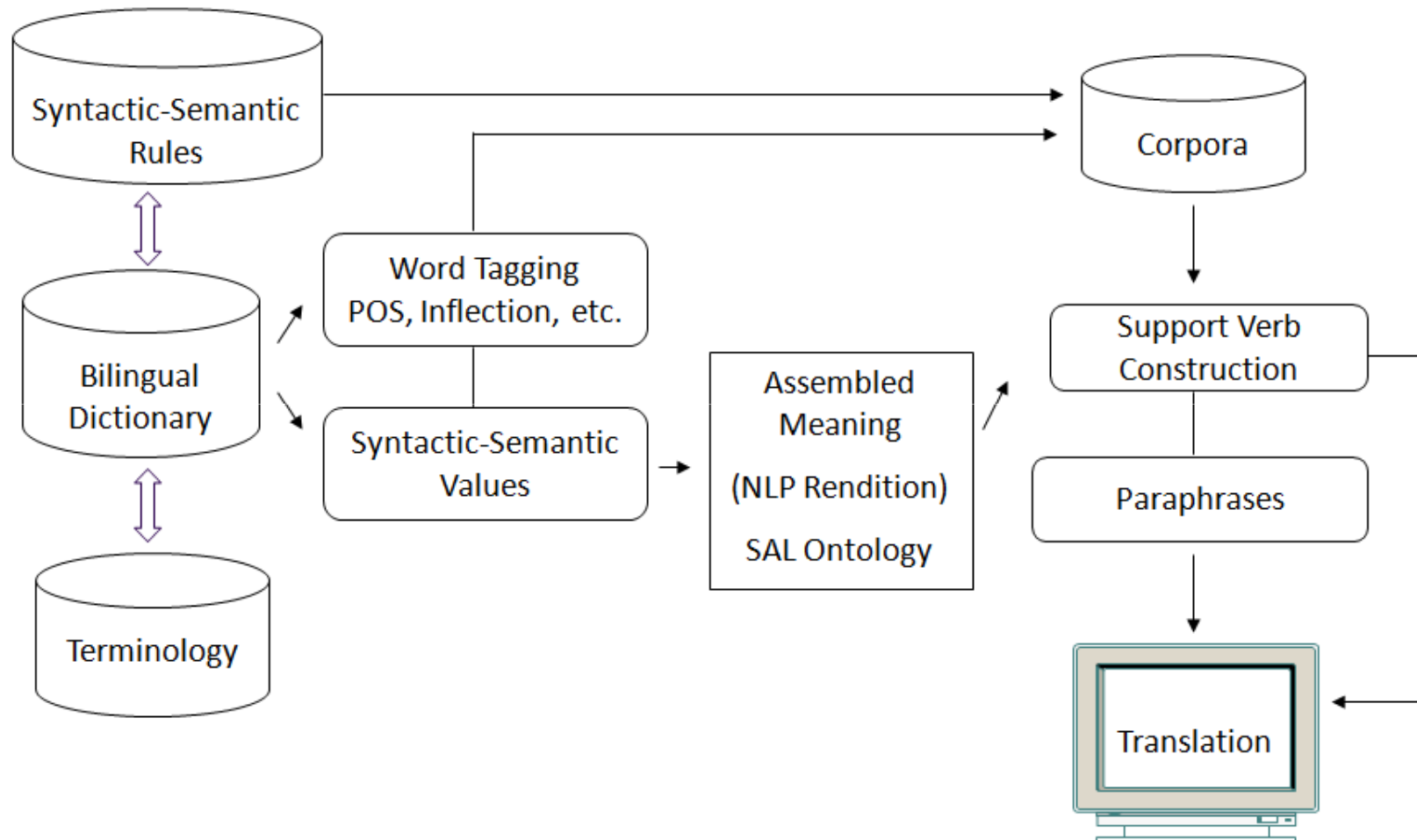
# APLICAÇÃO

## PARAMT: UM PARAFRASEADOR PARA A TA

a fazer um estágio para	dar aulas de/tutor	religião, mas não se impor
a fazer um estágio para	dar aulas de/lecture	religião, mas não se impor
a fazer um estágio para	dar aulas de/teach	religião, mas não se impor
ra que Edward pudesse	dar inicio /start	à carreira de clínico geral.
ra que Edward pudesse	dar inicio /initiate	à carreira de clínico geral.
ra que Edward pudesse	dar inicio /commence	à carreira de clínico geral.
ra que Edward pudesse	dar inicio /begin	à carreira de clínico geral.
os no Natal.» -- Estou a	dar aulas para/tutor	me livrar da tropa -- expli
os no Natal.» -- Estou a	dar aulas para/lecture	me livrar da tropa -- expli
Bernard aproveitou para	dar uma volta /walk	pela cozinha. Ainda há-de
Mas quem é que vai	dar um subsidio/subsidize	capaz para estudar os Oc
Mas quem é que vai	dar um subsidio/subsidise	capaz para estudar os Oc
a dizer-te que pode	dar a impressão de/seem	que não confio em ti
into que, se puder, devo	dar uma ajuda /help	ao Council. Um deles aba
into que, se puder, devo	dar uma ajuda /aid	ao Council. Um deles aba
into que, se puder, devo	dar uma ajuda /assist	ao Council. Um deles aba
ção de Persse, começa a	dar exemplos/exemplify	: O Espelho da Sinceridad
: irracional impedia-a de	dar voz /hear	ao medo. Ainda não eram
: irracional impedia-a de	dar voz /listen	ao medo. Ainda não eram
e sentia-se capaz de	dar um murro em/punch	quem tentasse detê-lo. M
ssel. Gostávamos de lhe	dar uma palavrinha/speak	. -- Então vamos dar um a
avrinha. -- Então vamos	dar um aperto de/shake	mão. «Deixar que o seu

Reconhecimento e parafraseamento bilingue de construções com verbos suporte (construção com verbo suporte em Português/ verbo lexical correspondente em Inglês)

# TRADUÇÃO AUTOMÁTICA COM PARÁFRASES



# BIBLIOGRAFIA SOBRE PARÁFRASES

- Anabela Barreiro. Make it simple with paraphrases: Automated paraphrasing for authoring aids and machine translation. PhD dissertation. Faculdade de Letras da Universidade do Porto. Porto, Portugal. December 2008. [http://www.linguateca.pt/Repositorio/AB-Thesis\\_030409.pdf](http://www.linguateca.pt/Repositorio/AB-Thesis_030409.pdf)  
[http://www.linguateca.pt/Repositorio/Phd\\_Thesis\\_Defense\\_Anabela\\_Barreiro\\_May25.pdf](http://www.linguateca.pt/Repositorio/Phd_Thesis_Defense_Anabela_Barreiro_May25.pdf)
- Nitin Madnani, Philip Resnik, Bonnie J. Dorr & Richard Schwartz, [Are Multiple Reference Translations Necessary? Investigating the Value of Paraphrased Reference Translations in Parameter Optimization](#), Proceedings of the Eighth Conference of the Association for Machine Translation in the Americas (AMTA 2008).
- Chris Callison-Burch, 2007. [Paraphrasing and Translation](#). PhD Thesis, University of Edinburgh. [\[slides\]](#)
- Chris Callison-Burch, Philipp Koehn, and Miles Osborne. 2006. [Improved Statistical Machine Translation Using Paraphrases](#). In Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (NAACL-2006), pp. 17–24. [\[slides\]](#)
- David Kauchak and Regina Barzilay. 2006. Paraphrasing for automatic evaluation. In proceedings of the Human Language Technology Conference of the NAACL, Main Conference, pp. 455–462, June.
- Liang Zhou, Chin-Yew Lin, and Eduard Hovy. 2006. Re-evaluating machine translation results with paraphrase support. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, pages 77–84, July. <http://acl.ldc.upenn.edu/W/W06/W06-1610.pdf>
- Grazia Russo-Lassner, Jimmy Lin, and Philip Resnik. [A Paraphrase-Based Approach to Machine Translation Evaluation](#). Technical Report LAMP-TR-125/CS-TR-4754/UMIACS-TR-2005-57, University of Maryland, College Park, August 2005.
- M. Dras. 1999. Tree Adjoining Grammar and the Reluctant Paraphrasing of Text. PhD Thesis, Macquarie University.

# BIBLIOGRAFIA SOBRE PARÁFRASES

## Introdução

Tese de [Regina Barzilay, 2003]

Tese de [Anabela Barreiro, 2008]

## Actas de Conferências e Workshops

ACL 2003 Workshop on Paraphrasing

3rd International Workshop on Paraphrasing (IWP2005)

ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (2007)

## Bibliografia

Página pessoal de Fujita

<http://paraphrasing.org/index.html>

## Outras teses

[Mark Dras, 1999]

[Florence Duclaye, 2003]

[Atsushi Fujita, 2005]

[Callison-Burch, 2007]

[Nakov, 2007]

## Outros trabalhos interessantes

Dekang Lin, Lillian Lee, Kevin Knight, Satoshi Sekine, Hua Wu, Ido Dagan, Oren Glickman, Yusuke Shinyama, Kozareva, Dolan, Brockett, Ravichandran, Hovy, Pantel, Szpektor, Bhagat, Hagiwara, ...

[adapt. Rodriguez, 2009]

# BIBLIOGRAFIA SOBRE LINGUAGEM CONTROLADA

- Rascu, Ecaterina, [A Controlled Language Approach to Text Optimisation in Technical Documentation](#). 8th Conference on Natural Language Processing. KONVENS 2006. Konstanz, 4-7 Oktober 2006. pp. 107-114.
- Maryline Hernandez and Ecaterina Rascu. 2004. Checking and Correcting Technical Documents. In Séverine Vienney and Mounira Bioud, editors, *Correction automatique: bilan et perspectives*, number 29 in BULAG, pp. 69–84. Presses universitaires de Franche-Comté.
- Teruko Mitamura, Kathryn Baker, Eric Nyberg, and David Svoboda. 2003. Diagnostics for Interactive Controlled Language Checking. In *Proceedings of EAMT/CLAW–2003*, pp. 87–94.
- Michael Carl, Johann Haller, Christoph Horschmann, Dieter Maas, and Jörg Schütz. 2002. The TETRIS Terminology Tool. *TAL, Structuration de terminologie*, 43(1):31–35.
- Haller, J.: 2000, 'MULTIDOC. [Authoring Aids for Multilingual Technical Documentation](#)'. In: *Proceedings of the 1st Congress of Specialized Translation*. Barcelona, pp. 143-147.
- Ursula Reuther and Antje Schmidt-Wigger. 2000. Designing a Multi-Purpose CL Application. In *Proceedings of CLAW–2000*, Seattle, WA.
- Antje Schmidt-Wigger. 1998. Grammar and Style Checking for German. In *Proceedings of CLAW–1998*, pages 76–86, Pittsburg, PA.
- Michael Carl and Antje Schmidt-Wigger. 1998. Shallow Postmorphological Processing with KURD. In *Proceedings of NeMLaP3/ CoNLL98*, Sydney.
- Arendse Bernth. *EasyEnglish: A tool for improving document quality*. In. Fifth Conference on Applied Natural Language Processing, pages 159–165, Washington, DC, USA, 1997. <http://acl.ldc.upenn.edu/A/A97/A97-1024.pdf>

# BIBLIOGRAFIA SOBRE LINGUAGEM CONTROLADA PARA TA

- Muegge, Uwe (2007). "[Controlled language: the next big thing in translation?](#)". ClientSide News Magazine (ClientSide Publications) **7** (7): 21–24.
- Michael Carl. 2001. Example-based Decomposition, Generalization and Refinement for Machine Translation. Ph.D. thesis, Universität des Saarlandes.
- Teruko Mitamura and Eric Nyberg. 2001. Automatic Rewriting for Controlled Language Translation. In Proceedings of NLPRS–2001. Workshop on Automatic Paraphrasing: Theory and Application.
- *Arendse Bernth, Claudia Gdaniec, MTranslatability, Machine Translation, v.16 n.3, p.175-218, September 2001.*  
<http://www.springerlink.com/content/p55758h300573742/>

# CONTROLO DA QUALIDADE LINGUÍSTICA E PARÁFRASES

-

## AJUDANTES DA TRADUÇÃO AUTOMÁTICA

Anabela Barreiro

LINGUATECA/FCCN

CLUP

**Aspectos do Processamento da Linguagem Natural em português**

**Ferramentas e recursos existentes > investigação futura**