

Plan de Trabajo Doctoral

**Traducción automática basada en conocimiento  
lingüístico y simplificación textual**

*Alumna:*

Ivana Romina ALTAMIRANO

*Supervisora:*

Dra. Paula ESTRELLA

FaMAF, 7 de septiembre de 2010

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Motivación del plan de trabajo</b>	<b>4</b>
2.1. Incorporación de recursos lingüísticos y simplificación textual . . . . .	4
2.2. Combinación de modelos de traducción . . . . .	5
2.3. Evaluación de la TA . . . . .	5
<b>3. Aplicabilidad</b>	<b>6</b>
<b>4. Objetivos</b>	<b>7</b>
<b>5. Plan de Actividades</b>	<b>8</b>
<b>Bibliografía</b>	<b>8</b>

# 1. Introducción

La **Traducción Automática** (TA, del Inglés *Machine Translation* - MT) se encarga de estudiar diversos métodos para convertir un texto dado en un lenguaje origen (LO) en otro texto semánticamente equivalente en una lengua destino (LM), con la menor intervención humana posible. Esta tarea es una de las más complejas del área del Procesamiento de Lenguaje Natural (PLN) dado que implica analizar el texto origen, tratando con ambigüedades a niveles léxico - estructural - funcional - pragmático, y generar texto en el idioma destino.

La idea de traducir lenguajes naturales de manera mecanizada, surge de comunicaciones informales entre algunos investigadores estadounidenses y comienza a concretarse a principio de los años 50 [Hutchins, 1995]. Inicialmente se propuso trabajar en aspectos lingüísticos complejos que permitieran la traducción automatizada de textos sin ningún tipo de intervención humana y con resultados de calidad similar a las traducciones realizadas por humanos. Este enfoque, conocido como *Fully Automatic High Quality MT*, fue descartado con el paso del tiempo como consecuencia de las sucesivas demostraciones de la no factibilidad de tal tarea.

Sin embargo, las dos décadas que duró este proceso permitieron el refinamiento de los métodos propuestos para la TA, evolucionando desde la **“traducción directa”** (simple reemplazo de palabras origen por palabras destino, utilizando sólo un diccionario bilingüe) a la **“traducción por transferencia”** (utilizando análisis del texto origen más profundo, representaciones formales como interlingua y/o reglas gramaticales de transferencia para pasar de una representación origen a una destino). Esta primer etapa de investigación en el área también dejó como resultado una serie importante de sistemas y proyectos que demostraron la efectividad y factibilidad de la TA para algunas tareas específicas, por ejemplo, la creación de reportes del tiempo [Chandioux, 1976], la traducción de resúmenes escritos en lenguajes controlados [Ducrot, 1973], o la traducción de textos técnicos, por ejemplo, matemáticos [Loh, 1976].

A mediados de los '80 surge otro paradigma de TA llamado **“basado en ejemplos”**, que propone traducir por analogía comparando el texto dado con un corpus paralelo y combinando en la salida aquellos segmentos relacionados, de acuerdo a alguna estrategia de extracción, selección y ranking [Nagao, 1984]. Una década más tarde, se propone utilizar métodos estadísticos en el proceso de TA, utilizados originalmente en el área de la teoría de la información [Shannon, 1948], en lo que se conoce como **“traducción estadística”** [Brown et al., 1993] y que constituye actualmente el paradigma de investigación más popular.

En la actualidad existen una gran variedad de sistemas y de herramientas (comerciales, de código libre, estadísticos, basados en reglas, en ejemplos, etc) para la creación de sistemas de TA de base que luego se adaptan o mejoran para cumplir un propósito particular, ya sea para organizaciones gubernamentales, empresas privadas, profesionales del lenguaje free-lance o investigadores. Sin embargo, ninguno de estos sistemas muestra excepcionales habilidades para traducir cualquier tipo de texto entre cualquier par de idiomas, sino que los sistemas que mejor performance muestran (los de más “alta calidad de traducción”)

son aquellos a los que se les agregan **recursos lingüísticos** específicos necesarios para un caso particular de uso.

Por ejemplo, tanto el sistema comercial **Systran** (basado en reglas desarrollado manualmente desde mediados de 1970 [Senellart et al., 2003]) como Language Weaver (primer sistema comercial estadístico [Benjamin et al., 2003]) son utilizados por numerosos clientes pero luego de su implantación se realiza una etapa de adaptación para cada uso particular. Para los sistemas desarrollados como consecuencia de investigaciones, este proceso de adaptación también se reconoce como beneficioso para mejorar la calidad de las traducciones, dado que actuando sobre un dominio restringido se disminuye las posibilidades de errores debido a la complejidad de las traducciones (ver por ejemplo [Koehn and Schroeder, 2007; Bouillon et al., 2007]).

Otra estrategia que se propone beneficiosa para la TA es la **simplificación textual** [Kring, 2001], es decir, aplicar un conjunto de reglas que permitan transformar un texto origen en uno equivalente pero de menor complejidad lingüística, por ejemplo, partiendo oraciones largas en una serie de oraciones más cortas, reemplazando ciertas palabras por sinónimos, detectando y eliminando redundancias, etc. La simplificación de texto podría contribuir a la generación de traducciones más acertadas lo que, a su vez, contribuiría a la mejor comprensión del texto generado; por ejemplo, se ha demostrado que las traducciones largas que presentan errores, son más difíciles de entender y evaluar que aquellas que son cortas y erróneas. Aunque la simplificación textual ha sido aplicada en diversas situaciones, no tenemos conocimiento de que haya trabajos que la apliquen directamente al proceso de la TA, por lo tanto la primera línea de investigación propuesta es **explorar la simplificación textual aplicada a la TA** e intentar demostrar su eficacia o rechazar la misma.

La segunda línea de investigación que se propone en este plan, es la **combinación de modelos** para la TA, es decir, se traduce el mismo texto origen con distintos modelos y se elige la mejor combinación de salida. Esta forma de explotar las mejores cualidades de cada modelo se ha vuelto muy popular a través de competencias que permiten la comparación de sistemas en un marco común (utilizando los mismos recursos, traduciendo los mismos textos y siendo evaluados con las mismas métricas) [Callison-Burch et al., 2009]. Sin embargo, no resulta trivial la combinación de distintas traducciones ya que esto implica “entenderlas” o analizarlas de alguna manera, y esta etapa se realiza generalmente utilizando métricas que describan la calidad de cada una de las TA en cuestión. Es por ello, que otra área de investigación muy ligada al presente plan de trabajo, es la **evaluación de la TA**, la cual interviene durante el proceso de combinación de sistemas (al elegir aquellos segmentos de mejor calidad) y al final del proceso de traducción, cuando se aplican diversas métricas para determinar la calidad total del sistema.

## 2. Motivación del plan de trabajo

### 2.1. Incorporación de recursos lingüísticos y simplificación textual

En la sección anterior mencionamos que la incorporación de recursos lingüísticos incrementa la calidad de los sistemas de TA, especialmente si son estadísticos o basados en ejemplos, dado que este tipo de sistemas sólo se basa en un corpus paralelo alineado sin ninguna anotación lingüística adicional y genera la salida de acuerdo a ciertos indicadores estadísticos calculados sobre el corpus de entrenamiento. El conocimiento lingüístico es particularmente útil cuando se traduce hacia un idioma morfológicamente rico, como es el caso del Español.

Dentro del conocimiento que se podría agregar a modelos estadísticos, podemos nombrar los **marcos de sub-categorización verbal** (recordar que los verbos en idiomas como el español son las palabras que más variación presentan) que permitirían evitar errores en la traducción si conseguimos guiar la misma de acuerdo a los patrones encontrados.

Otro recurso lingüístico que puede ser usado son los **n-gramas de información morfológica** del lenguaje destino, es decir, podrían estudiarse cómo resolver inconsistencias en las traducciones a través de las posibles combinaciones de palabras de acuerdo a su función sintáctica; por ejemplo, el paquete de traducción Moses [Hoang and Koehn, 2003] permite agregar distintas dimensiones (llamados factores) que describan distintos aspectos del lenguaje origen, luego uno de ellos puede ser el factor morfológico y se podría usar esta información durante el proceso de traducción así como también durante la generación de la salida para corregir errores de concordancia; un ejemplo muy sencillo es controlar la concordancia artículo-sustantivo impidiendo que aparezcan cosas del estilo <artículo-singular-masculino sustantivo-plural-femenino> en la salida.

A un nivel de análisis más profundo, más allá del análisis sintáctico y morfológico, hay indicios de que la incorporación de **aspectos semánticos** a la TA puede mejorar sustancialmente los resultados, como se muestra en [Fung et al., 2001]. Esta hipótesis sólo ha sido explorada para el par de idiomas Chino a Inglés y para un corpus particular para el cual se cuenta con las anotaciones semánticas necesarias. Para validar estas conclusiones preliminares, es imperativo comprobarlas en diversas situaciones, que incluyan distintos tipos de idiomas y dominios; en particular, nuestro interés es estudiar la incorporación de roles semánticos a la TA con español como idioma destino.

Adicionalmente, se explorará la simplificación de texto, que es la tarea de hacer a los textos más fácil de leer y entender, pero manteniendo el contenido y la semántica, esta tarea puede incluir estrategias de resumen, simplificación léxica y sintáctica, reemplazo de alguna estructura por otra más simple, reorden y parafrasis. Cabe mencionar que la simplificación incluye las etapas de análisis, transformación y generación.

La simplificación de texto se ha usado mayormente en educación, tanto para la educación de idiomas extranjeros [Petersen, 2007], como en la ayuda a las personas con problemas de afasia, dislexia u otros [Carroll et al., 1996], también se ha usado en extracción, recuperación de información y como asistencia a la lectura [Inui et al., 2003]. Por lo

tanto, nuestra intención es explorar las técnicas de simplificación disponibles en las áreas mencionadas anteriormente y aplicarlas a la traducción automática, y así poder decidir si alguna de ellas mejora la traducción.

La principal limitación del uso de recursos lingüísticos adicionales es que la mayoría están disponibles libremente sólo para el Inglés o, si existen para otros idiomas, suelen presentarse en versiones más acotadas, con lo cual la investigación se ha focalizado mayoritariamente en pares de idiomas que incluyen al Inglés como idioma origen o destino (ver por ejemplo un resumen en [Kuhn and Isabelle, 2009]). En nuestro caso, el idioma de interés es el Español por el impacto regional de estudiar nuestro propio idioma y por la disponibilidad de hablantes nativos, necesaria para la evaluación de la TA. En este sentido, una parte importante de nuestro trabajo se dedicará a estudiar los recursos disponibles para el español, su aplicabilidad y, si fuese necesario, su extensión y/o creación de nuevos recursos.

## 2.2. Combinación de modelos de traducción

La utilización de un único modelo de traducción (ya sea basado en reglas, en ejemplos o estadístico) se ha demostrado eficaz en una serie de situaciones que incluyen dominios restringidos o la especialización del sistema de TA a un dominio particular pero desde hace algún tiempo se reconoce también que fallan en dominios abiertos [Nirenburg and Frederking, 1994]. Es por ello, que una alternativa es la implementación de modelos híbridos que permitan explotar mejor las características especiales de cada modelo, lo cual ha indicado una mejoría en los resultados; por ejemplo, los modelos estadísticos pueden tener mayor cobertura fuera del dominio de entrenamiento (favoreciendo el recall) mientras que los sistemas basados en reglas o ejemplos favorecen la precisión de las traducciones.

En este aspecto, se propone explorar las formas tradicionales de combinar sistemas pero agregando una etapa de pre-edición que conste del análisis sintáctico de la oración de entrada (particularmente su descomposición en sub-frases) y una etapa de post-edición que consista en la verificación y corrección de aspectos estéticos de las traducciones, los cuales tienen gran influencia sobre la percepción que tienen los usuarios sobre la calidad de la TA. La principal motivación para trabajar en etapas es que la traducción de unidades más simples (frases más cortas, posiblemente simplificadas) facilitaría la re-combinación de traducciones candidatas (proceso) aumentando la calidad en la etapa de generación.

## 2.3. Evaluación de la TA

La evaluación de sistemas de generación de lenguaje natural es una de las más difíciles dentro del área del PLN, dado que una idea puede expresarse de muchas formas, todas ellas correctas. En general, determinar la calidad de una frase generada no puede hacerse de manera simple y directa, por ejemplo, comparando el resultado del sistema con un patrón (en Inglés *gold standard*). Para la evaluación de la calidad de las traducciones se han propuesto diversas metodologías que pueden clasificarse en *directas* e *indirectas*. Una metodología directa aplica alguna métrica al texto generado por un sistema y a su vez pueden clasificarse

en “humanas” (por ejemplo [Brown et al., 1999; Doyon et al., 1998; Vilar et al., 2007]) y “automáticas” (por ejemplo [Babych, 2004; Banerjee and Lavie, 2005; Lin and Och, 2004]). Una metodología indirecta, en cambio, evalúa la performance del sistema a través de la utilización del texto generado para realizar alguna tarea (por ejemplo [Fuji and Isahara, 2001; Laoudi et al., 2006; White and Taylor, 1998]). El área de la evaluación es muy activa y existe una gran diversidad de métricas tanto humanas como automáticas, de hecho, hay competencias (llamadas campañas de evaluación) dedicadas exclusivamente a la meta-evaluación de la TA, donde, por ejemplo en 2010 se han propuesto 26 nuevas métricas sólo en ese evento. Sin embargo, ninguna de las metodologías propuestas hasta el momento se ha demostrado eficaz de manera de ser aceptada en la comunidad.

Por otro lado, para determinar la calidad de los prototipos obtenidos (incluyendo calidad de traducción, funcionalidad, etc.) se propone estudiar la creación de un modelo de la calidad según los estándares de evaluación de software ISO/IEC 9126 [ISO/IEC, 2001] y 14528 [ISO/IEC, 1999], los cuales fueron exitosamente aplicados al dominio de la TA, como lo muestra la herramienta FEMTI (Framework for the Evaluation of Machine Translation) [Estrella et al., 8/2009]. FEMTI intenta guiar a los evaluadores hacia la creación de planes de evaluación parametrizables que incluyen diversos aspectos del sistema a evaluar y ofrece un conjunto de métricas (humanas y automáticas) a aplicarse según la característica de calidad elegida por el evaluador.

La identificación de métricas relevantes se puede realizar usando distintos métodos: basándose en experiencias previas [Hajdinjak and Mihelic, 2006], realizando encuestas o especificaciones de requerimientos (como en [Lecoeuche et al., 1998]) o bien recolectando estos datos a través de experimentos llamados “mago de Oz” (del Inglés wizard of Oz) donde el usuario interactúa con un prototipo de sistema posiblemente incompleto o reducido en funcionalidades y un humano (el “mago”) detrás de la interface responde como si lo hiciera el sistema [Dahlbäck et al., 1998; Starlander, 2007].

### 3. Aplicabilidad

Si bien el sistema de traducción propuesto en esta tesis mejorará el estado del arte de la traducción automática, según la disponibilidad de recursos se explorará la posibilidad de integrar las mejoras propuestas a uno de los siguientes:

- Un prototipo de sistema que provea el servicio de lectura de traducciones de artículos técnicos para alumnos de Famaf, o generador de material de estudio específico para algún tema de computación. La idea sería traducir papers y usar un sintetizador de voz para producir audio. En esta aplicación podríamos explorar las posibilidades de conseguir recursos en portugués, ruso, francés.
- Un tutor de aprendizaje de segundas lenguas integrado en un entorno virtual, en este caso particular el sistema permitiría ampliar el dominio del sistema deduciendo lo que el usuario quiso decir y mostrando sugerencias para que el usuario mejore la frase y así aprenda como se dicen las cosas en el idioma en cuestión.

- Una herramienta libre de pre y post edición que permita visualizar recursos lingüísticos usados por el sistema de traducción, e integrando también memorias de traducción. Esta herramienta será de particular interés para profesionales dedicados a la traducción, localización y afines.

## 4. Objetivos

En esta sección detallaremos los objetivos relevantes, donde los puntos 1 al 4 corresponden a la primera etapa de esta tesis doctoral, cuyo calendario de actividades se describe en la Sección 5.

1. Relevamiento bibliográfico de las áreas involucradas inicialmente (simplificación textual, combinación de modelos de traducción, introducción de conocimiento a modelos estadísticos)
2. Realizar un relevamiento de los recursos disponibles para las tareas propuestas
  - a) Determinación de parámetros iniciales para el trabajo propuesto en función de los recursos disponibles; estos parámetros incluyen: seleccionar el/los pares de idiomas para la TA, dominio de los textos a utilizar y herramientas específicas a aplicar (parsers, taggers, tokenizers, diccionarios, etc).
3. Estudio y aplicación de las herramientas potencialmente necesarias
  - a) Estudiar los motores de traducción automática disponibles open source; un relevamiento preliminar indica que hasta el momento los paquetes de interés son: Moses [Koehn et al., 2007], Cunei [Phillips, 2010], Joshua [Li et al., 2009] y Apertium [Ramírez-Sánchez et al., 2006]
  - b) Estudiar las herramientas disponibles para simplificación de texto e implicación textual, estudiar los recursos lingüísticos disponibles para los idiomas que se estudiarán, y la posibilidad de extenderlos para otros idiomas
4. Documentación y diseminación de resultados: esta etapa puede realizarse, por ejemplo, a través de comunicaciones en encuentros regionales/nacionales/internacionales, a través de presentaciones de progreso en las reuniones del grupo de PLN, etc.
5. Propuesta de estrategias de simplificación de texto
6. Propuesta de cómo y a qué modelos integrar los recursos lingüísticos seleccionados
7. Implementación de prototipos que integren las propuestas de los puntos anteriores
8. Evaluación de los prototipos creados: algunas metodologías a utilizar incluyen la creación de un modelo de la calidad basados en los estándares ISO/IEC mencionados

anteriormente, estudios comparativos con otras aplicaciones similares via la aplicación de métricas humanas/automáticas, la participación en competiciones pertinentes como el *Workshop for Statistical Machine Translation* o la evaluación basada en la tarea realizada por un grupo de sujetos.

9. Documentación y disseminación de resultados; preparación de herramientas generadas para su publicación.

## 5. Plan de Actividades

A continuación se sumarian las tareas a desarrollar en los próximos doce meses comprendidos entre el 1 de Septiembre 2010 y el 31 de Agosto 2011.

Tarea	1	2	3	4	5	6	7	8	9	10	11	12
1. Relevamiento y estudio de material bibliográfico	×	×	×	×								
2. Relevamiento de los recursos disponibles				×	×	×						
3. Estudio y aplicación de las herramientas seleccionadas						×	×	×	×			
4. Estudiar los modelos de TA seleccionados									×	×	×	×
5. Documentación y disseminación de resultados											×	×

## Bibliografía

- B. Babych. Weighted N-Gram Model for Evaluating Machine Translation Output. In *Proceedings of the 7th Colloquium on Computational Linguistics in the UK (CLUK)*, Birmingham, UK, 2004.
- S. Banerjee and A. Lavie. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. In *Proceedings of the Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization at 43rd ACL*, Ann Arbor, Michigan, 2005.
- B. Benjamin, L. Gerber, K. Knight, and D. Marcu. Language Weaver: The Next Generation of machine Translation. In *Proceedings of the IX Machine Translation Summit*, pages 23–27, New Orleans, Louisiana, 2003.
- P. Bouillon, S. Halimi, M. Rayner, and B. A. Hockey. Adapting a Medical Speech to Speech Translation System (MedSLT) to Arabic. In *Semitic '07: Proceedings of the 2007 Workshop on Computational Approaches to Semitic Languages*, pages 41–48, Morristown, NJ, USA, 2007. Association for Computational Linguistics.
- J. Brown, A. Yamashiro, and E. Ogane. Three Strategies for Tailoring Cloze Tests in Secondary EFL. Technical Report TUJ Working Papers in Applied Linguistics 14, Temple University, Japan, 1999.

- P. Brown, V. Della Pietra, S. Della Pietra, and R. Mercer. The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(2):263–311, 1993.
- C. Callison-Burch, P. Koehn, C. Monz, and J. Schroeder. Findings of the 2009 Workshop on Statistical Machine Translation. In *Proceedings of the Fourth Workshop on Statistical Machine Translation*, pages 1–28, Athens, Greece, 2009. Association for Computational Linguistics.
- J. Carroll, G. Minnen, Y. Canning, S. Devlin, and J. Tait. Practical Simplification of English Newspaper Text to Assist Aphasic Readers. 1996.
- J. Chandiooux. METEO, an Operational System for the Translation of Public Weather Forecasts. *FBIS Seminar on Machine Translation, American Journal of Computational Linguistics*, pages 27–36, 1976.
- N. Dahlbäck, A. Jönsson, and L. Ahrenberg. Wizard of oz studies—why and how. pages 610–619, 1998.
- J. Doyon, K. B. Taylor, and J. S. White. The DARPA MT Evaluation Methodology: Past and Present. In *AMTA Conference*, Philadelphia, PA, USA, 1998.
- J. M. Ducrot. Le système TITUS II. *A.N.R.T.Information et Documentation*, 4:3–40, 1973.
- P. Estrella, A. Popescu-Belis, and K. M. The femti guidelines for contextual mt evaluation: principles and tools. In W. Daelemans and V. Hoste, editors, *Evaluation of Translation Technology*. *Linguistica Antverpiensia*, 8/2009.
- M. Fuji and H. Isahara. Evaluation Method for Determining Groups of Users Who Find MT Useful. In B. Maegaard, editor, *MT Summit VIII*, pages 103–108, Santiago de Compostela, Galicia, Spain, 2001. <http://www.eamt.org/summitVIII/papers.html>.
- P. Fung, Z. Wu, Y. Yang, and D. Wu. Automatic Learning of Chinese English Semantic Structure Mapping. 27(3):373–408, 2001.
- M. Hajdinjak and Mihelic. The paradise evaluation framework: Issues and findings. *Computational Linguistics*, 32(2):263–272, 2006. ISSN 0891-2017.
- H. Hoang and P. Koehn. Moses: Statistical Machine Translation System. Master’s thesis, University of Edinburgh, 2003.
- W. J. Hutchins. Machine translation: A brief history. In *Concise History of the Language Sciences: from the Sumerians to the Cognitivists*, pages 431–445. Pergamon Press, 1995.
- K. Inui, A. Fujita, T. Takahashi, R. Iida, and T. Iwakura. Text Simplification for Reading Assistance: A Project Note. In *In Proceedings of the 2nd International Workshop on Paraphrasing: Paraphrase Acquisition and Applications*, pages 9–16, 2003.
- ISO/IEC. *14598-1:1999 (E) – Information Technology – Software Product Evaluation – Part 1: General Overview*. International Organization for Standardization / International Electrotechnical Commission, Geneva, 1999.

- ISO/IEC. *9126: Information Technology – Software Product Evaluation / Quality Characteristics and Guidelines for Their Use*. International Organization for Standardization / International Electrotechnical Commission, Geneva, 2001.
- P. Koehn and J. Schroeder. Experiments in Domain Adaptation for Statistical Machine Translation. In *StatMT '07: Proceedings of the Second Workshop on Statistical Machine Translation*, pages 224–227, Morristown, NJ, USA, 2007. Association for Computational Linguistics.
- P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst. Moses: Open Source Toolkit for Statistical Machine Translation. In *ACL*. The Association for Computer Linguistics, 2007.
- H. Kring. *Repairing texts: empirical investigations of machine translation post-editing processes*. Kent State University Press, 2001.
- R. Kuhn and P. Isabelle. MT: the Current Research Landscape. <http://www.mt-archive.info/MTS-2009-Kuhn-ppt.pdf>, 2009. Keynote Presentation at The twelfth Machine Translation Summit, Canada.
- J. Laoudi, C. Tate, and C. Voss. Task-Based MT Evaluation: from Who/When/Where Extraction to Event Understanding. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*, pages 2048–2053, Genoa, Italy, 2006.
- R. Lecoecue, C. Mellish, and D. S. Robertson. A framework for requirements elicitation through mixed-initiative dialogue. In *ICRE '98: Proceedings of the 3rd International Conference on Requirements Engineering*, page 190, Washington, DC, USA, 1998. IEEE Computer Society. ISBN 0-8186-8356-2.
- Z. Li, C. Callison-Burch, C. Dyer, J. Ganitkevitch, S. Khudanpur, L. Schwartz, W. N. G. Thornton, J. Weese, and O. F. Zaidan. Demonstration of Joshua: An Open Source Toolkit for Parsing-based Machine Translation. In *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*, pages 25–28, Suntec, Singapore, August 2009. Association for Computational Linguistics.
- C. Lin and F. Och. Automatic Evaluation of Machine Translation Quality Using Longest Common Subsequence and Skip-Bigram Statistics. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 605–612, Barcelona, Spain, 2004.
- S.-C. Loh. CULT: Chinese University Language Translator. *FBIS Seminar on Machine Translation, AJCL*, pages 46–50, 1976.
- M. Nagao. A Framework of a Mechanical Translation between Japanese and English by Analogy Principle. In *Proc. of the international NATO symposium on Artificial and human intelligence*, pages 173–180, New York, NY, USA, 1984. Elsevier North-Holland, Inc.
- S. Nirenburg and R. Frederking. Toward multi-engine machine translation. In *HLT '94:*

- Proceedings of the workshop on Human Language Technology*, pages 147–151, Morristown, NJ, USA, 1994. Association for Computational Linguistics.
- S. E. Petersen. Natural Language Processing Tools for Reading Level Assessment and Text Simplification for Bilingual Education. 2007.
- A. Phillips. The Cunei Machine Translation Platform for WMT’10. In *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR*, pages 155–160, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- G. Ramírez-Sánchez, F. Sánchez-Martínez, S. Ortiz-Rojas, J. A. Pérez-Ortiz, and M. L. Forcada. Opentrad apertium open-source machine translation system: an opportunity for business and research. In *Proceeding of Translating and the Computer 28 Conference*, November 2006.
- J. Senellart, C. Boitet, and L. Romary. SYSTRAN New Generation: The XML Translation Workflow. In *Proc. of the Ninth Machine Translation Summit*, pages 338–345, New Orleans, LA, EUA, 2003.
- C. E. Shannon. A mathematical theory of communication. *Bell system technical*, 27, 1948.
- M. Starlander. Using a Wizard of Oz as a Baseline to Determine which System Architecture is the Best for a Spoken Language Translation System. In *Proceedings of the 16th Nordic Conference of Computational Linguistics (Nodalida)*, pages 161–164, 2007.
- D. Vilar, G. Leusch, H. Ney, and R. Banchs. Human Evaluation of Machine Translation Through Binary System Comparisons. In *Proceedings of the 2nd Workshop on Statistical Machine Translation*, pages 96–103, Prague, Czech Republic, 2007.
- J. White and K. Taylor. A Task-Oriented Evaluation Metric for Machine Translation. In *Proceedings of the 1st International Conference on Language Resources and Evaluation (LREC)*, pages 21–25, Granada, Spain, 1998.