

Evaluación automática de un sistema híbrido de predicción de palabras y expansiones

Sira E. Palazuelos Cagigas

José L. Martín Sánchez

Universidad de Alcalá

Escuela Politécnica Superior. Campus

Universitario s/n. 28805. Alcalá de Henares.

{sira, jlmartin}@depeca.uah.es

Javier Macías Guarasa

Grupo de Tecnología del Habla

Universidad Politécnica de Madrid

Ciudad Universitaria s/n. 28040. Madrid.

macias@die.upm.es

Resumen: La predicción de palabras es uno de los sistemas más utilizados para ayudar a la escritura a personas con problemas físicos y/o lingüísticos. Últimamente la predicción de palabras se complementa con otras estrategias para mejorar su rendimiento como la expansión de abreviaturas o predicción de frases. En este artículo se presenta un sistema híbrido, de predicción de palabras y predicción de expansiones (es decir, se expande la abreviatura incluso antes de acabar de escribirla). En este sistema se permite al usuario abreviar o no cada palabra, y reducir la carga cognitiva requerida para su utilización, ya que no se necesita memorizar abreviaturas fijas para cada palabra. La eficiencia del sistema se evalúa en base al porcentaje de pulsaciones que ahorra con respecto a la escritura del mismo texto sin ayuda, mostrándose resultados de la predicción de palabras y de expansiones por separado y de la combinación de ambos.

Palabras clave: Predicción de palabras, expansión de abreviaturas, predicción de expansiones, modelado del lenguaje, ayudas a la escritura y comunicación para personas con discapacidad.

Abstract: Word prediction is one of the most commonly used systems to help to write people with physical and/or linguistic disabilities. In the newest systems, word prediction is complemented with other strategies to improve its performance, such as abbreviation expansion or phrase prediction. In this paper, a hybrid system with prediction of words and expansions is presented. Expansion prediction consists in expanding the abbreviation even before the user finishes writing it. This system allows the user to abbreviate or not a word, and reduces the cognitive load required for its use because it is not necessary to remember a fixed abbreviation for each word. The parameter used to evaluate the efficiency of the system is the percentage of keystrokes saved with respect to writing the text without help, and we include results of the word prediction, the expansion prediction and the combination of both.

Keywords: Word prediction, abbreviation expansion, expansions prediction, language modeling, technical aids for writing and communication for people with disabilities.

1 Introducción

La predicción de palabras consiste en ofrecer al usuario posibles terminaciones al fragmento de palabra que haya escrito, de forma que, si se predice la palabra que busca, seleccione la predicción y no necesite acabar de escribir la palabra. Es una de las técnicas más utilizadas para ayudar a escribir texto y comunicarse a personas con distintas discapacidades.

Inicialmente su objetivo era reducir el número de pulsaciones (y con ello el tiempo)

que los usuarios con discapacidad física necesitaban para escribir un texto, pero estudios posteriores han demostrado que: no siempre se produce realmente una aceleración en la escritura (al menos en las etapas de uso iniciales), que los usuarios con problemas físicos valoran más la reducción de esfuerzo físico necesario para producir el texto y que los usuarios con problemas lingüísticos también podían utilizarlo para producir textos más correctos (Magnuson y Hunnicutt, 2002).

Para generar la lista de palabras predichas se utilizan diferentes técnicas de modelado del lenguaje como las descritas en (Allen, 1994). Numerosos sistemas usan modelos basados en n-gramas para generar las palabras predichas, como, por ejemplo, el descrito en (Leshner, Moulton y Higginbotham, 1999), que muestra unos resultados de 54,7% para trigramas con listas de predicción de 10 palabras para inglés. (Carlberger et al, 1997) presenta un sistema de predicción para sueco, inglés, danés, noruego, francés, ruso y español basado en ngramas y en información de las últimas palabras utilizadas (*recency*). En versiones previas, como la descrita en (Hunnicut, 1989) utilizaban también información semántica en el proceso de predicción. En la versión siguiente han incorporado modelos de Markov para palabras y categorías (Hunnicut y Carlberger, 2001) presentando un ahorro de pulsaciones para sueco de un 46% con una lista de 5 palabras predichas. En (Garay-Vitoria y Gonzalez-Abascal 1997) se presenta un sistema basado en un chart parser, que más tarde han adaptado a las características particulares del vasco, idioma con un alto grado de flexión en (Garay-Vitoria, Abascal y Gardeazabal, 2002). En este último artículo proponen utilizar gramáticas con reglas que describan la sucesión de categorías que forman una categoría compuesta, y la predicción basada en morfemas con posibilidad de aceptación de palabras completas. El resultado que consiguen para vasco con listas de 5 palabras predichas es aproximadamente del 43%.

En la actualidad la predicción de palabras está siendo complementada con otras técnicas como la expansión de abreviaturas (Leshner y Moulton, 2005), (Willis et al., 2002) y (Willis, Pain y Trewin, 2005), y la predicción de frases (Väyrynen, Noponen y Seppänen, 2007).

Los algoritmos de expansión de abreviaturas se pueden dividir en fijos y flexibles. En su gran mayoría desarrollan mecanismos de desabreviación automática y aceptan cierto margen de error como (Willis et al., 2002), (Willis et al., 2005). La diferencia fundamental de los sistemas del mercado y el descrito en (Palazuelos et al., 2006), que es evaluado en este artículo, es que los algoritmos de expansión de abreviaturas revisados anteriormente expanden una abreviatura después de que ésta haya sido escrita completamente, mientras que en este artículo se

habla de predicción flexible de expansiones: se proponen expansiones al fragmento escrito de la abreviatura en curso (aunque no se haya acabado de escribir). Otra diferencia es que las palabras, en este trabajo, no tienen asignadas abreviaturas fijas, sino que cada persona puede abreviarlas como desee mientras siga ciertas reglas de compresión. También difiere de los anteriores en que propone un sistema de expansión directamente supervisado por el usuario, es decir, se predicen las expansiones a la vez que se escribe el texto y se muestran al usuario las candidatas para que él elija la deseada y la inserte, obteniendo así un texto final totalmente correcto, sin margen de error.

La estructura del artículo es la siguiente: en primer lugar se describe brevemente la arquitectura del sistema de predicción de palabras y expansiones. A continuación se muestran los resultados de ambos sistemas de predicción por separado y combinados. Finalmente, se exponen las conclusiones.

2 Descripción del sistema de predicción de palabras y expansiones

El algoritmo de predicción (tanto de palabras como de abreviaturas) consta básicamente de tres bloques que son explicados en detalle en (Palazuelos, 2001) y (Palazuelos et al, 2006):

- Diccionarios.
- Modulo de predicción.
- Interfaz de usuario.

Los **diccionarios** contienen palabras y unidades multipalabra y toda la información (gramatical y probabilística) que necesitan los métodos de predicción. El sistema contiene un diccionario general para castellano de más de 150.000 entradas, y diccionarios temáticos y personales adaptables al usuario y a la temática del texto que se está escribiendo, que aumentan la probabilidad de predicción de las palabras que ya se han escrito en el texto o que han aparecido en textos sobre el mismo tema. Además, también se han entrenado de forma automática diccionarios para otros idiomas, como el inglés o el portugués.

Los **métodos de predicción**, a partir del texto escrito por el usuario, proponen restricciones que deben cumplir las palabras siguientes (categoría gramatical y su probabilidad, concordancias, etc.). Los métodos

de predicción disponibles están basados en secuencias de hasta 6 palabras (n-gramas), hasta 3 categorías (n-POS) y un analizador basado en una gramática independiente del contexto, cuya potencia ha sido aumentada de forma importante para soportar: gestión de probabilidades de reglas, ambigüedad (gramatical) de las palabras, posibilidad de que en la regla haya elementos (terminales o no terminales) opcionales, posibilidad de que los símbolos no terminales sean tanto categorías gramaticales como significantes o lemas (imponiendo las reglas de concordancia de rasgos adecuadas), posibilidad de prohibir un determinado significante o lema en una posición determinada de una regla, y un potente sistema de tratamiento de rasgos, que permite tanto controlar la concordancia entre los distintos símbolos (terminales y no terminales), como imponer o prohibir rasgos en cualquier símbolo de una regla.

La **interfaz de usuario** se encarga de recoger el texto que está siendo escrito, recibir las restricciones de los métodos de predicción a partir de ese texto, obtener de los diccionarios el listado de palabras que cumplen dichas restricciones y mostrarle las más probables al usuario como listado de palabras predichas.

La

Figura 1 muestra un teclado virtual que incluye los algoritmos de predicción de palabras y de expansiones. La predicción, además, está incluida en otros sistemas de ayuda a personas con discapacidad como el sistema de comunicación *Comunicador*, aplicación de acceso gráfico a mensajes descrita en (Palazuelos 2005), o PredWin, editor de texto con acceso por barrido muy utilizado en España por la comunidad de personas con graves discapacidades físicas (Palazuelos 2001).

←	Esp	↵	⊗	Supr	Signos	⇐	⇓	⇒	⇑		Tecla
e	a	i	⊗	r	t	b	Tab	May	Shift		Teclado
o	s	d	⊗	u	q	v	Ctrl	Alt	Foco		Telecomunicación
n	l	m	⊗	g	y	z	Inicio	Fin	Menu		Teleférico
c	p	f	⊗	h	j	k	ReP	AvP	Salir		Telefonista
'	,	.	⊗	x	ñ	w	F1-12				Teléfono
											Teletexto

Figura 1: Ventana de edición de la aplicación *Comunicador*, incluyendo la lista de palabras y expansiones predichas tras escribir "Tel"

A partir de la información de los diccionarios y los métodos de predicción, el algoritmo de predicción de palabras mostrará al usuario las palabras más probables que

comiencen exactamente por el fragmento escrito de la palabra en curso, y cumplan las restricciones impuestas por los métodos de predicción.

El algoritmo de predicción de expansiones propuesto tiene un funcionamiento similar al de predicción de palabras, pero, a la hora de comparar el fragmento escrito de la palabra en curso con las palabras del diccionario, aplica una serie de reglas de expansión, tales como:

- Aplicación de los **heurísticos** más frecuentes, por ejemplo, fonéticos o de sustitución (x=por).
- Búsqueda en diccionarios por similitud de cadena teniendo en cuenta que puede haber **letras eliminadas**.
- Expansión **fija** por medio de **tablas** de pares abreviatura-expansión.
- Se está estudiando la inclusión de **aprendizaje automático de abreviaturas**, aunque el hecho de que el sistema sea flexible hace que el aprendizaje se reduzca a los heurísticos y las tablas fijas.

El algoritmo de expansión es explicado en detalle en (Palazuelos et al., 2006).

3 Evaluación automática del sistema

La importancia de la predicción radica, no sólo en su capacidad para acelerar la tasa escritura o la comunicación, sino también en el aumento en la calidad del texto generado por una persona, y la disminución del esfuerzo, tanto físico como cognitivo, necesario para escribirlo. Estos y otros resultados se muestran en (Magnuson y Hunnicutt, 2002) en un estudio a largo plazo, en el que se pudo constatar tanto la reducción en el número de pulsaciones como la aceleración en la escritura a lo largo de los 13 meses de duración del experimento.

La disminución en el esfuerzo cognitivo, (especialmente en personas con dislexia, que cometen demasiadas faltas de ortografía o con cualquier otro problema que provoque que generen textos de baja calidad) es muy difícilmente evaluable de forma automática, y se deja la valoración a expertos que puedan comprobar el aumento en la calidad de los textos generados. Este aumento en la calidad suele conllevar un aumento en la cantidad, ya que los usuarios se sienten más capaces de

escribir textos correctos, y se produce una realimentación positiva en el proceso.

En cuanto a la evaluación de la disminución del esfuerzo físico que se produce por la realización de las pulsaciones necesarias para escribir el texto, la métrica que mejor lo refleja es el porcentaje de ahorro de pulsaciones con respecto a la escritura sin ayuda de predicción. Este parámetro sí puede ser evaluado de forma automática.

Hemos de considerar que, además de los muchos factores que influyen en la eficacia de la predicción (tanto el idioma, como la configuración del sistema de predicción o de la propia interfaz donde esté instalado (Palazuelos et al., 1999) como subjetivos por preferencias del usuario), si el sistema de predicción no es capaz de predecir la palabra adecuada y reducir el número de pulsaciones necesarias, los demás factores serán irrelevantes (Trnka et al., 2005). Por eso es tan importante realizar una evaluación automática del porcentaje de pulsaciones ahorrado.

Para realizar una evaluación automática del sistema, se utiliza un **modelo de usuario** que simula a una persona escribiendo texto y eligiendo siempre las predicciones correctas cuando se muestran (usuario perfecto). Se toma el texto carácter a carácter y se llama al algoritmo de predicción que hace una propuesta de las posibles palabras predichas después de escribir cada letra. Si alguna de estas palabras se corresponde con la que se está intentando escribir, el sistema la elige, contabilizándola como palabra predicha correctamente y acumulando el ahorro de pulsaciones que produce.

La selección de los textos de entrenamiento y prueba constituye uno de los aspectos más importantes a la hora de realizar la evaluación de cualquier técnica de procesamiento del lenguaje natural (PLN) y se realiza teniendo en cuenta aspectos explicados en (Palazuelos, 2001). En esta serie de experimentos se deseaba evaluar la calidad en la escritura de texto (uso habitual de *PredWin*, editor de texto, y del teclado virtual, dos de las aplicaciones donde está incluida la predicción), no de conversación (como *Comunicador*). Se utilizó un texto de prueba resultado de la combinación de varios cuentos, con una longitud de 2000 palabras, teniendo en cuenta que los usuarios de estos sistemas, con graves discapacidades

físicas, normalmente no escriben textos muy grandes en cada sesión.

Como referencia se contabiliza la cantidad de pulsaciones necesaria para escribir el texto sin ningún algoritmo de ayuda cuyos datos se muestran en la Tabla 1.

Nombre texto de prueba	“Cuentos variados”
Número de palabras	2000
Num. pulsaciones para escribirlo sin ayuda	11969

Tabla 1: Datos sobre el texto de prueba

3.1 Evaluación automática del algoritmo de predicción de palabras

En el primer experimento se utiliza predicción de palabras, con 5 candidatas en la lista de predicción, sin ningún tipo de ayuda gramatical, solamente la información estadística contenida en el diccionario general (de más de 150.000 entradas) obteniéndose los resultados que se muestran en la Tabla 2.

Nombre texto de prueba	“Cuentos variados”
Número de palabras	2000
Núm. pulsaciones con predicción de palabras sin ayuda gramatical	7937
% ahorro de pulsaciones	33,68%

Tabla 2: Resultados de la predicción de palabras sin ayuda gramatical

Posteriormente se introduce el análisis gramatical basado en secuencias de categorías gramaticales (POS, parts of speech), *bipos* y *tripos* (Allen, 1994).

Nombre texto de prueba	“Cuentos variados”
Número de palabras	2000
Num. pulsaciones para escribirlo con predicción de palabras utilizando tripos	7701
% ahorro de pulsaciones	35,65%

Tabla 3: Resultados de la predicción de palabras usando tripos

Según puede verse en la Tabla 3, el ahorro de pulsaciones mejora en un 1,97 % con respecto al anterior.

Si, además, incorporamos la utilización de los n-gramas y el diccionario del texto en curso, el ahorro es mucho mayor como podemos observar en la Tabla 4.

Nombre texto de prueba	“Cuentos variados”
Número de palabras	2000
Num. pulsaciones con predicción de palabras utilizando tripos, n-gramas y diccionario en curso	7243
% ahorro de pulsaciones	39,48%

Tabla 4: Predicción de palabras con tripos, n-gramas y el diccionario de texto en curso

Los resultados de la Tabla 4 muestran que al utilizar los n-gramas, además de los bipos, tripos y el diccionario de texto en curso, se produce una mejora de un 3,83% respecto a los resultados obtenidos aplicando solo tripos y de un 5,8% si no se aplica ningún mecanismo de ayuda gramatical durante la predicción.

3.2 Evaluación automática del algoritmo de predicción de expansiones

Los parámetros de evaluación son los mismos que para la predicción de palabras. La evaluación automática es realizada con un modelo de usuario más complejo que el de la predicción de palabras, ya que debemos considerar que escribe texto abreviado. Por esto necesitamos utilizar dos ficheros: el texto con el que deseamos realizar la evaluación y su versión abreviada.

Debido a la dificultad para disponer de corpus paralelos abreviados y sin abreviar, ha sido necesario implementar un proceso para comprimir automáticamente los ficheros de prueba, aplicando las siguientes técnicas de compresión (que intentan imitar en lo posible las estrategias de compresión habituales de los usuarios de teléfono móvil):

- Las palabras más frecuentes se comprimen aplicando heurísticos (fonéticos, etc.)
- Se eliminan las letras cuyo porcentaje de aparición en el texto supere un 2%

- Se incluye una estrategia de compresión fija con tabla, es decir, si una palabra o secuencia de palabras está en dicha tabla, se sustituye directamente por la abreviatura asociada.
- Las palabras menos frecuentes se mantienen sin comprimir, ya que la probabilidad de que el sistema las descomprima es reducida. Debemos considerar que deseamos un texto totalmente libre de error, es decir, que si la abreviatura se acaba de escribir y no se ha descomprimido, el sistema simulará un retroceso, y reescribirá la palabra sin comprimir (sumando las pulsaciones necesarias para realizar todo este proceso). Si al comprimir el texto dejamos sin abreviar las palabras menos frecuentes este proceso se elimina, o al menos se reduce, penalizando menos los resultados. Hemos de tener en cuenta que los usuarios también comprimen poco/nada las palabras poco frecuentes, para evitar que quien lea el mensaje pueda pensar que la abreviatura se corresponde con otra palabra más frecuente.

Este archivo comprimido es el utilizado para realizar la evaluación automática. No obstante se realizarán futuras evaluaciones con usuarios reales donde se espera conseguir mejores resultados, teniendo en cuenta que la inteligencia del usuario hará que utilice la estrategia óptima en base al funcionamiento de la expansión.

Además, se ha incluido en la evaluación otra circunstancia que también puede darse en casos reales: si el usuario está utilizando el sistema para comunicarse, necesita velocidad y que el texto sea comprensible, aunque no sea perfecto, y premiará la rapidez a la corrección total. En este caso puede que no corrija las abreviaturas que no se expandan si el texto resultante se puede entender sin dificultad. Se ha introducido esta posibilidad en el sistema, y en los experimentos se proporcionan también resultados considerando que puede haber **margen de error** (abreviaturas sin descomprimir).

En esta serie de experimentos se ha utilizado el mismo texto de prueba que en los experimentos anteriores. A continuación se evalúa el ahorro de pulsaciones aplicando

únicamente el algoritmo de predicción de expansiones, haciendo uso de tripos y n-gramas aplicadas a los diccionarios general y personal. Los resultados obtenidos se muestran en la Tabla 5.

Nombre texto de prueba	“Cuentos variados”
Num. pulsaciones con predicción de expansiones sin error	6461
Ahorro de pulsaciones	46,01%

Tabla 5: Predicción de expansiones sin error

Si no se tienen en cuenta los retrocesos, es decir, si se admite un cierto porcentaje de abreviaturas sin descomprimir (margen de error), los resultados obtenidos se muestran en la Tabla 6.

Nombre texto de prueba	“Cuentos variados”
Num. pulsaciones con predicción de expansiones con error	6415
Ahorro de pulsaciones	46,40%

Tabla 6: Predicción de expansiones con error

Estos resultados se obtuvieron con un porcentaje de error de un 0,6%, muy bajo respecto a otros sistemas revisados como el descrito en (Shieber y Baker, 2003) que presenta un 3%.

Según puede apreciarse, el sistema de predicción de expansiones sin error obtiene un ahorro de pulsaciones de un 46,01% y con error se ahorra un 46,40%, implicando un incremento del 0,39% en el ahorro de pulsaciones. Además, se puede observar que las mejoras con respecto a la predicción de palabras (Tabla 4) son de 6,53% y 6,92% respectivamente.

3.3 Eficacia de la combinación de los algoritmos de predicción de palabras y expansiones

Es posible configurar el modelo de usuario para que se pueda introducir texto normal y abreviado, y el programa es capaz de generar una lista de posibles palabras predichas combinando las propuestas de los algoritmos de predicción de palabras y expansiones.

En esta sección se evalúa la eficacia de la combinación de estos dos algoritmos con respecto a la utilización de cada uno de ellos

por separado. Se comparan los resultados dando prioridad a cada uno de los algoritmos de predicción. Esto quiere decir que en cada experimento se puede decidir cual de los dos algoritmos será el primero en realizar la propuesta de palabras predichas, y si una vez rellena esta lista de posibles palabras, esta no está completa, se llama al otro algoritmo de predicción para que la complete con su propuesta. Es decir, por cada letra que introduzca el usuario, se mostrará una lista de cinco posibles palabras procedentes del algoritmo prioritario o de los dos.

En la Tabla 7 se muestran los resultados obtenidos al darle prioridad al algoritmo de predicción de expansiones frente al de predicción de palabras. Según puede apreciarse, los resultados mejoran un 3% respecto a la aplicación del algoritmo de predicción de expansiones por sí solo, sin tener en cuenta errores.

Nombre texto de prueba	“Cuentos variados”
Num. pulsaciones ambos algoritmos prioridad expansión sin error	6094
Ahorro de pulsaciones	49,08%

Tabla 7: Combinación de algoritmos dando prioridad a la predicción de expansiones

Por otro lado, si se da prioridad a la predicción de palabras, los resultados obtenidos se muestran en la Tabla 8.

Nombre texto de prueba	“Cuentos variados”
Num. pulsaciones ambos algoritmos prioridad predicción sin error	5606
Ahorro de pulsaciones	53,16%

Tabla 8: Combinación algoritmos dando prioridad a la predicción de palabras

En este caso el ahorro de pulsaciones mejora más de un 4% con respecto a los resultados obtenidos dando prioridad a la predicción de expansiones.

4 Conclusiones

En este artículo se evalúa la eficacia de los algoritmos de predicción de expansiones y

palabras que se utilizan en varios sistemas de ayuda a la escritura y comunicación para personas con discapacidad. Para realizar la evaluación automática de los algoritmos presentados se ha diseñado un modelo de usuario capaz de simular la entrada de texto en cada caso.

En primer lugar se exponen los resultados obtenidos aplicando sólo el método de predicción de palabras. La introducción de información gramatical permite que no se presenten al usuario predicciones gramaticalmente incorrectas, y esto produce una mejora en los resultados obtenidos de un 1,97%, además de una mejora subjetiva en la calidad apreciada por el usuario. En el siguiente experimento, además de los tripos, se utilizan los n-gramas y el diccionario personal, logrando un ahorro de pulsaciones de un 39,48% que equivale a una mejora de un 3,83% respecto al método anterior.

Posteriormente se evalúan los resultados considerando que el usuario escribe texto abreviado y se aplica el algoritmo de **predicción de expansiones**. También se considera si se admite un **margen de error** en el texto o no (el porcentaje de error obtenido no supera el 0,6% en ningún caso). El ahorro de pulsaciones obtenido sin error fue de un 46,01% mejorando los resultados obtenidos con los algoritmos de predicción de palabras en un 6,9%.

La **combinación** de los dos algoritmos de predicción permite que el usuario introduzca texto abreviado o texto normal, y produce los mejores resultados cuando se da prioridad a la predicción de palabras con un ahorro de pulsaciones en el orden de un 53,16% libre de error, lo cual supera en un 4% al algoritmo que da prioridad a la predicción de expansiones, en más de un 7% al mejor de los algoritmos de predicción de expansiones y en casi un 14% al mejor algoritmo de predicción de palabras.

Por último, debemos considerar que la introducción de estos algoritmos en el sistema de ayuda a la escritura y/o comunicación no sólo ofrece ventajas cuantitativas en base al ahorro de pulsaciones, sino que también da **flexibilidad** al usuario a la hora de abreviar, permitiendo que comprima cada vez de una manera diferente y no necesite recordar la abreviatura asignada a cada palabra, por lo tanto, **reduce la carga cognitiva** que supondría memorizarlas.

Bibliografía

- Allen, J. 1994. "Natural language Understanding". Benjamin/Cummings Publishing Company Inc 2ª Ed.
- Carlberger A., Carlberger J., Magnuson T. Hunnicutt S., Palazuelos-Cagigas S., Aguilera Navarro S. 1997. Profet, a new generation of word prediction: An evaluation study. Proceedings, ACL Workshop on Natural language processing for communication aids, 23–28, Madrid.
- Garay-Vitoria N. and Gonzalez-Abascal. J. 1997. Intelligent word prediction to enhance text input rate (a syntactic analysis-based word prediction aid for people with severe motor and speech disability). In Proceedings of the Annual International Conference on Intelligent User Interfaces, 241–244.
- Garay-Vitoria N. Abascal J., Gardeazabal L. 2002. "Evaluation of Prediction Methods Applied to an Inflected Language". Lecture Notes In Computer Science; Vol. 2448. Proceedings of the 5th International Conference on Text, Speech and Dialogue Pages: 389 – 396. ISBN:3-540-44129-8.
- Hunnicutt, S. 1989. "Using Syntactic and Semantic Information in a Word Prediction Aid". Proc. Europ. Conf. Speech Commun. Paris, France. September 1989, vol. 1. páginas: 191-193.
- Hunnicutt S., Carlberger J. 2001. "Improving Word prediction using Markov models and heuristic methods". Augmentative and Alternative Communication, Volume 17, Issue 4 December, pages 255 – 264.
- Leshner, G.W., Moulton, B.J., Higginbotham, D.J. (1999). Effects of ngram order and training text size on word prediction. Proceedings of the RESNA'99 Annual Conference, 52-54, Arlington, VA: RESNA Press.
- Leshner G., Moulton B., 2005. "An introduction to the theoretical limits of abbreviation expansion performance". 28 Annual RESNA Conference Proceedings. <http://www.dynavotech.com/files/research/LeMo05.pdf>
- Magnuson T., Hunnicutt S., 2002. "Measuring the effectiveness of Word prediction: The advantage of long-term use". Speech, Music

- and Hearing, KTH, Estocolmo, Suecia. TMH-QPSR. Volumen 43: 57-67.
- Palazuelos S. E., Aguilera S., Rodrigo J. L., Godino. J., Martín J. 1999. Considerations on the Automatic Evaluation of word prediction systems. *Augmentative and Alternative Communication: New Directions in Research and Practice*. Pags: 92-104. Whurr Publishers. Londres.
- Palazuelos Cagigas S. 2001. “Aportación a la predicción de palabras en castellano y su integración en sistemas de ayuda a personas con discapacidad física”. Tesis Doctoral.
- Palazuelos Cagigas S. E., Martín Sánchez J. L., Arenas García J., Godino Llorente J. I., Aguilera Navarro S. 2001. “Communication strategies using PredWin for people with disabilities”. *Conference and Workshop on Assistive Technology for Vision and Hearing Impaired*. Castelvecchio Pascoli, Italia. Agosto.
- Palazuelos Cagigas S. E., Martín Sánchez J. L., Domínguez Olalla L. M. 2005. “Graphic Communicator with Optimum Message Access for Switch Users”. *Assistive technology: from virtuality to reality*. Pags: 207-211. ISBN: 1-58603-543-6. ISSN: 1383-813X. Ed. IOS Press (A. Pruski y H. Knops).
- Palazuelos Cagigas S. E., Martín Sánchez J. L., Hierrezuelo Sabatela L., Macías Guarasa J. 2006. “Design and evaluation of a versatile architecture for a multilingual word prediction system”. LNCS (Lecture Notes in Computer Science) 4061. *Computers Helping People with Special Needs*. Springer-Verlag. Editores: Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, Arthur Karshmer. ISBN: 3-540-36020-4. Páginas 894-901.
- Trnka, K., Yarrington, D., McCoy, K., Pennington, C., 2005. “The Keystroke Savings Limit in Word Prediction for AAC”. <http://hdl.handle.net/123456789/149>.
- Shieber S., Baker E. 2003. “Abbreviated Text Unput”, IUI’03, Miami, Florida, USA. ACM 1-58113-586-6/03/0001. 12-15 Enero 2003. <http://www.iuiconf.org/03pdf/2003-001-0064.pdf>
- Väyrynen P., Noponen K., Seppänen T. 2007. “Analysing performance in a word prediction system with multiple prediction methods”. *Computer, Speech & Language* Volume 21. Issue 3. Páginas 479-491. Julio.
- Willis T., Pain H., Trewin S., Clark S. 2002. “Informing Flexible Abbreviation Expansion for users with motor disabilities”. *Lecture Notes In Computer Science; Vol. 2398 Proceedings of the 8th International Conference on Computers Helping People with Special Needs*. Páginas: 251 – 258. ISBN: 3-540-43904-8.
- Willis T., Pain H., Trewin S. 2005. “A Probabilistic Flexible Abbreviation Expansion System for Users With Motor Disabilities”. *School of Informatics, University of Edinburgh*.