

Casandra: original sistema de apoyo a la escritura

S. G. Luvoni¹, P. D. Agüero¹, J. C. Tulli¹, E. L. Gonzalez¹, A.J. Uriz¹, J.M. Garín¹

Universidad Nacional de Mar del Plata, Facultad de Ingeniería: Mar del Plata, Argentina

¹sluvoni@fi.mdp.edu.ar, pdaguero@fi.mdp.edu.ar, jctulli@fi.mdp.edu.ar, elgonzal@fi.mdp.edu.ar, ajuriz@fi.mdp.edu.ar, jmgarin@fi.mdp.edu.ar

Resumen. Actualmente existe una gran necesidad de sistemas de apoyo a la escritura de textos para personas discapacitadas que pueden ser utilizados con toda la variedad de software disponible en el mercado.

En este artículo presentamos el sistema Casandra soporte de la escritura de textos, que permite interactuar con software existentes en el mercado. Su apariencia es una ventana que permanece siempre visible en la pantalla ofreciendo un pequeño número de palabras, una de las cuales podría ser la deseada por el usuario. De esta manera este se evitara escribir la totalidad de las letras de la palabra.

Palabra Clave: Predicción, Discapacidad, Escritura, Ordenador.

1.-Introducción

Las aplicaciones de predicción de palabras constituyen una importante ayuda a las personas discapacitadas en la escritura por ordenador. El objetivo de una aplicación de la predicción, es hallar la palabra que desea escribir el usuario con el fin de evitarle que deba ingresar la totalidad de letras de la misma. Por ejemplo, en el caso que el usuario desee escribir la palabra **hola**, el sistema intentará predecir con cada letra ingresada la palabra deseada.

En un primer momento, el usuario recibe un listado de posibles palabras cuando todavía no ha ingresado ninguna letra como se muestra en la primera columna de la figura 1. Luego de ingresada la letra H, el listado es diferente. Todas ellas comienzan con la letra H. Sin embargo, todavía la palabra "HOLA" no aparece y el usuario debe continuar ingresando letra por letra la palabra.

Finalmente, luego de ingresada la letra "O", la palabra "HOLA" aparece en segundo lugar del listado. En este momento, el usuario puede seleccionar la palabra y se evitará el ingreso de las letras "L", "A" y el espacio posterior a la palabra.

Listado inicial de palabras	Listado de palabras luego de "H"	Listado de palabras luego de "HO"
El	hacia	Hora
La	hasta	Hola
Los	he	Hoy
Las	ha	Hombre
De	han	Homenaje
En	Hemos	Horizonte

Tabla 1: Listado de palabras dependiendo de las letras ingresadas.

Este sistema de apoyo a la escritura es de amplia aplicabilidad tanto en discapacitados como en personas que deseen mejorar su velocidad de escritura. En el mercado existe software que ofrece esta funcionalidad dedicado a un grupo particular de usuarios con necesidades específicas: Profet [1], PredWin [2] (Figura 2), Fazly [3], Fasty [4], WordQ [5]. Casandra se diferencia de ellos en algunas características que permiten un mejor funcionamiento y que además son transparentes al usuario, tal como se mostrará en las secciones siguientes.

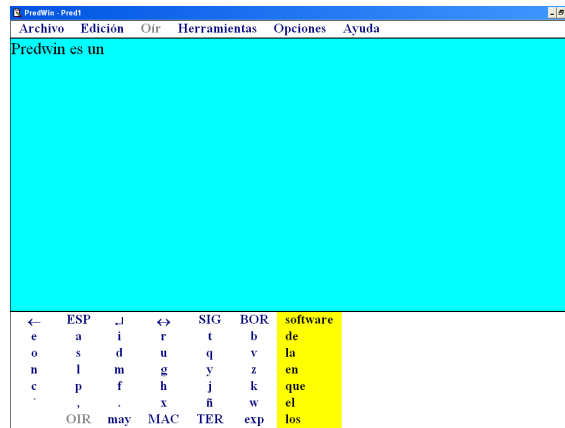


Figura 1: Ejemplo de pantalla de predicción del software PredWin.

El objetivo de nuestro trabajo es ofrecer un sistema de predicción de textos aplicable a cualquier software del mercado a través de una ventana que siempre se mantiene visible permitiendo la interacción del usuario con el sistema de predicción. De esta manera se puede seleccionar libremente la herramienta que se usará y también amplía el espectro de aplicaciones que pueden ser utilizadas: procesadores de texto, mensajería instantánea, navegación por Internet, enciclopedias, conversores de texto a voz, software profesional, etc.

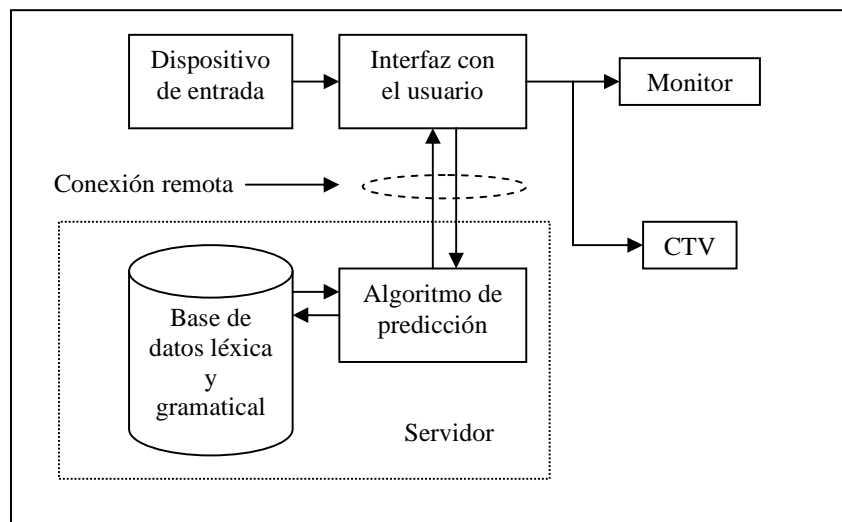


Figura 2: Esquema del sistema cliente/servidor Casandra para el apoyo a la escritura por predicción de palabras.

El esquema del sistema se puede ver en la Figura 2. Básicamente consiste en un dispositivo de entrada (teclado, pulsador, sensor de iris, electrooculografía, etc.). El predictor ofrece a través de una interfaz visual adaptada opciones de escritura para su selección por parte del usuario. El proceso de predicción se efectúa en un servidor remoto por Internet o red de área local, de este modo los clientes pueden tener ser adaptados a diferentes plataformas, desde computadoras portátiles, agendas inteligentes, teléfonos celulares y diferentes sistemas operativos. Finalmente, el sistema también puede contar con otro tipo de salidas, además de la visual a través del monitor, tales como convertidores de texto a voz (CTV) para la comunicación oral.

2.- Metodología

En general, la arquitectura del sistema de predicción de palabras **Cassandra** se puede observar en la Figura 4.

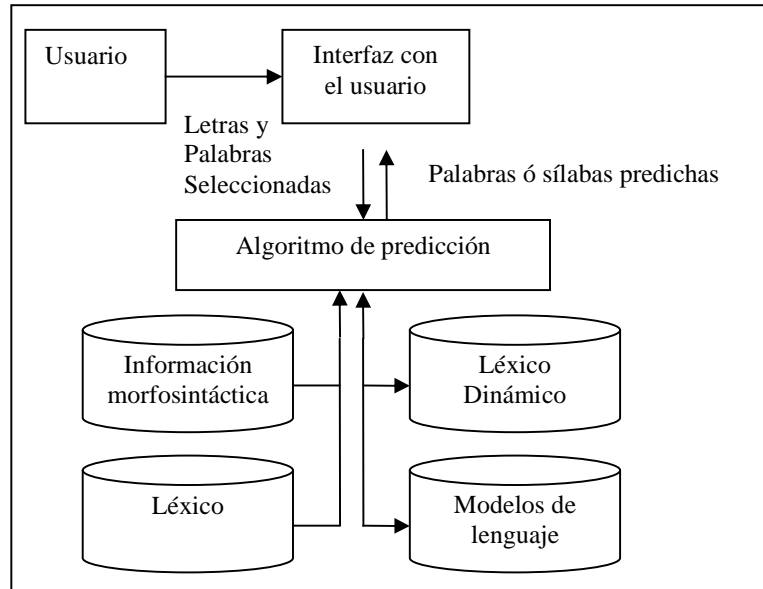


Figura 4: Arquitectura del sistema de predicción de palabras de Cassandra.

Allí se observan dos partes bien diferenciadas. Por un lado se encuentra la interfaz con el usuario que se adapta a las necesidades del mismo de acuerdo a su nivel de discapacidad para ingresar las letras y seleccionar las palabras.

Por otro lado tenemos el algoritmo de predicción que es totalmente independiente de la interfaz con el usuario.

El sistema predictor se basa tal como se ve en la figura en información de entrenamiento fija y una sección de entrenamiento dinámico que es el propio aporte del cliente que utiliza el sistema.

2.1 – La predicción silábica

Como novedad, Cassandra incorpora además un sistema de predicción de palabras en unidades de **silabas** por lo que, si las nuevas palabras que el cliente o usuario desean escribir no llegaran a encontrarse en la base de datos del servidor pero en palabras similares del mismo grupo o que posean asociaciones silábicas similares, el sistema presenta una lista de elección de las posibles sílabas que pueden aportar un ahorro en el tipeo de letras. El uso de estas unidades silábicas hace que la base de datos involucrada sea sensiblemente menor a la de palabras. Y como las palabras en el idioma castellano poseen reglas en la formación de sílabas, se puede deducir que solo es necesario un número pequeño de datos para formar cualquier palabra válida.

A modo de ejemplo: El cliente desea escribir "**muchachita**" pero en la base de datos sólo está registrada la palabra "**muchachitos**". Al tipear "m" aparece la siguiente lista de opciones a elegir: (las sílabas aparecen en negrita en esta demostración)

1.-muerte 2.-me 3.-mañana 4.-mano 5.-moralidad 6.-misma 7.-mi

Al tipear "mu":

1.-muerte 2.-muy 3.--**muy**- 4.-mucho 5.-mundo 6.-muchachitos 7.--**mun**- 8.--**muer**-

Al tipear "muc":

1.--**cho**= 2.--**cha**= 3.--**chas**= 4.--**chí**= 5.-mucho 6.-muchachito 7.-mucha 8.-muchas 9.-muchísimo 10.-muchachitos

Vemos que la opción 2 es la sílabas "cha" y que el cliente puede elegir y formar la palabra "mucha" donde ahorra un tipeo.

Siguiendo el ejemplo: Palabra "muchac":

1.--chi= 2.--chas= 3.-muchachitos 4.-muchachas

Luego la palabra queda "muchachi" al elegir la opción 1.

Al final, el usuario completaría con las dos letras que faltan para formar la palabra "muchachita" ahorrando dos tipesos: m-u-c-(ha)-c-(hi)-t-a.

A pesar que la palabra no estaba registrada en el sistema. Luego que este proceso, el módulo de léxico dinámico incorporara estas nuevas palabras haciendo que el proceso de predicción sea más eficiente. El procesamiento silábico permite obtener estadísticas de las sílabas en relación a su posición en la palabra (al principio, en la mitad o al final). Dependiendo del tipo de probabilidad, el sistema muestra en la lista de opciones de sílabas también la alternativa de elegir la sílaba con algún símbolo o carácter a continuación (por ejemplo, un espacio en blanco en el caso de las sílabas con altas probabilidades de pertenecer al grupo de sílabas de final de palabra). Este sistema permite excluir de la lista de selección información de poca relevancia, como grupos de letras o sílabas que no concuerden con las reglas de separación. Una consecuencia de este método es que el usuario puede tener ampliada la oferta de selección de datos y aumentar la eficiencia del ahorro de tipeo.

2.2 – Interfaz con el usuario

Esta interfaz realiza la selección de las letras, sílabas y palabras que se presentan al usuario. Existe una gran variedad de mecanismos de visualización adaptados a las necesidades de los mismos:

- **Listado de palabras y sílabas.** Este sistema está orientado a aquellos que pueden usar un teclado normal o adaptado y desean incrementar su velocidad de escritura mediante un algoritmo de predicción. En este caso se visualiza el listado de palabras que puede seleccionar el usuario mediante un conjunto de teclas predefinidas por el mismo. Estas teclas pueden ser creadas especialmente en el caso de un teclado adaptado o bien redefinir la funcionalidad de una tecla preexistente en el teclado disponible.
- **Listado de palabras, sílabas y letras con barrido vertical y horizontal.** En algunas ocasiones la persona no puede hacer uso de un teclado debido a su grado de discapacidad. En tales casos son necesarios mecanismos diferentes para permitir la selección de letras y palabras. Un sistema de barrido horizontal y luego vertical, permite la selección de letras y palabras en una matriz.
- **Electrooculografía y sensado del iris.** Mediante el uso de la vista y algún mecanismo de validación es posible seleccionar letras y palabras en una matriz. Este sistema es mucho más rápido que el anterior, ya que evita los retardos propios del barrido de la matriz [6].

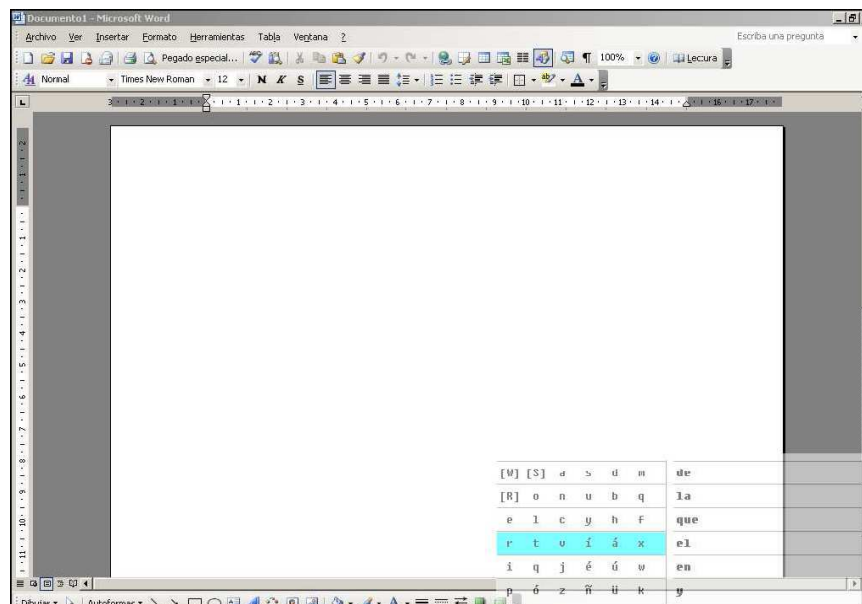


Figura 5: Un ejemplo de Interfaz utilizado para comunicarse con Casandra

Como Casandra en sí provee resultados de predicción según los datos que le son enviados, el tipo y diseño de la interfaz con el usuario será el mas conveniente según lo que considere el desarrollador y el usuario al cual esta dirigido.

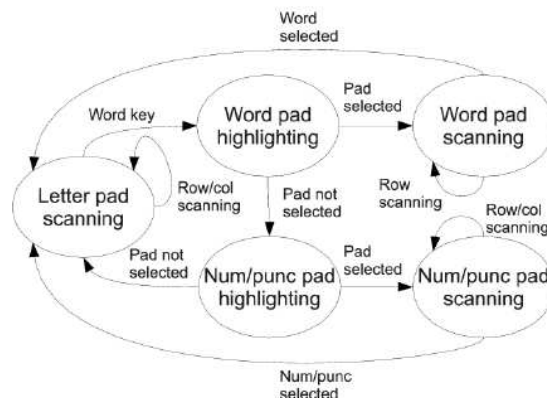


Figura 6: Diagrama de estado de la interfaz mostrada en la Figura 5.

2.3 – Métodos de predicción de palabras

Para la predicción de las palabras existe una gran variedad de algoritmos en la literatura. En nuestro caso hacemos uso de una combinación de técnicas y herramientas para esta tarea.

Léxico general.

El sistema más básico de predicción es el uso de un léxico con las palabras más frecuentes utilizadas en el idioma. Este sistema es idéntico al usado en los teléfonos móviles. El ingreso de sucesivas letras de la palabra deseada produce un filtrado de las posibles palabras del léxico, y por lo tanto, es posible que el listado luego de este filtrado contenga la palabra buscada.

Léxico adaptado

La afirmación anterior que “el léxico posee las palabras más frecuentes usadas en el idioma” es inexacto debido a que el vocabulario es altamente dependiente de diversos factores relacionados con el usuario y el tema del texto que se está escribiendo, entre otros.

Cada persona posee un vocabulario específico debido a su ubicación geográfica, trabajo, condición social, edad, interlocutor, etc. Todos estos aspectos constituyen características personales específicas que deben ser tenidas muy en cuenta y preservadas porque transmiten su identidad.

Por otra parte, el tema del texto a escribir define un vocabulario específico. Por ejemplo, es diferente el vocabulario para escribir una carta, que el utilizado en un servicio de mensajería instantánea (chat). En el caso de un chat se hace uso de estructuras gramaticales y secuencias de caracteres específicos, tales como los emoticones (por ejemplo: “:-)”) y expresiones especiales como: “mmmm!”.

Por estos motivos es necesario el uso de un léxico específico para tales dominios que se encuentre adaptado al usuario.

Información morfosintáctica

En muchos idiomas es necesario el uso de información morfosintáctica para aumentar la capacidad de predicción del algoritmo.

En el caso del español, los artículos, adjetivos y sustantivos concuerdan en género y número dentro de un sintagma nominal, mientras que el verbo concuerda en número y persona con el sujeto. Es fundamental tener en cuenta tales aspectos para reducir el número de las posibles palabras a visualizar por el usuario y evitarle presentar variantes completamente inadecuadas que puedan hacer perder su confianza en el sistema.

Modelo de lenguaje

Observando un texto es posible llegar a la conclusión de que una palabra depende de las escritas previamente. Los modelos estadísticos permiten hacer uso de esta observación para estimar la probabilidad de una palabra en función de las escritas anteriormente. Por ejemplo, la probabilidad de “blanca” aumenta luego de haber escrito “la casa” ($p(\text{blanca}|\text{la,casa})$) [7].

Estas probabilidades condicionales se pueden estimar usando textos de entrenamiento. Sin embargo, la estimación es una tarea difícil debido a la gran cantidad de posibles combinaciones de palabras a estimar. Las técnicas de suavizado permiten asignar probabilidades a combinaciones no observadas en los textos de entrenamiento. Y por otra parte, el uso de lemas permite reducir la cantidad de posibles combinaciones.

Reglas de separación o segmentación silábica

En el idioma español existen una serie de reglas para separar palabras en sílabas por lo que si estas palabras son "válidas", son divisibles en unidades silábicas. Como en las bases de datos de los sistemas de predicción no es posible guardar todas las palabras que existen, puede usarse las reglas de separación en sílabas para formar nuevas palabras válidas. Esto ayuda a reducir el tipeo de letras al presentar en la predicción sílabas de tres o mas letras que pueden ir a continuación de la palabra semicompleta que el cliente/usuario esta tratando de completar, aun cuando la palabra en si no se encuentre en la base de datos.

3.- Implementación y resultados.

Casandra dispone de una serie de características que permiten su mejor funcionamiento a través de una serie de algoritmos adaptativos, agregando la capacidad de predicción basada en segmentos silábicos.

El uso de léxicos adaptados permite incrementar la calidad en la predicción de palabras. Casandra hace uso de una estructura de léxicos generados dinámicamente a partir de textos ingresados por el usuario con anterioridad, de esta manera enriquece el diccionario de palabras de uso más frecuente. Estos léxicos se encuentran clasificados en un conjunto de clases mediante algoritmos automáticos de agrupamiento para su selección automática dependiendo del texto que se está ingresando actualmente. De esta manera, es transparente y evita la necesidad de selección por parte del usuario del léxico adaptado adecuado.

Por otra parte, en el momento de hacer la predicción se hace uso de una combinación de diferentes modelos de lenguaje, palabras, lemas, información morfosintáctica y la construcción de nuevas palabras basadas en las reglas silábicas del idioma español. Casandra posee un ajuste automático de los pesos de cada uno de los modelos de lenguaje al tomar la decisión sobre que palabras o sílabas se visualizarán en la lista. Esto permite adaptarse a circunstancias donde el vocabulario sea reducido y predecible, o bien asignar más peso a la estructura gramatical debido a la riqueza de las palabras usadas.

En la Figura 6 se muestra el porcentaje de ahorro de teclas mediante Casandra. Usando los métodos predictivos de estadísticas de n-gramas y diccionarios como base, vemos la influencia de usar la predicción silábica y el agregado del aprendizaje dinámico del vocabulario del usuario (léxico adaptado) en un escenario (*) donde el usuario repite estructuras y palabras luego de algunos párrafos.

- SS: Sin sílabas
- CSSP: Con sílabas sin Estadística Posicional
- CSPO: Estadística Posicional Inicial
- CSPM: Estadística Posicional Completa
- CSPL: CSPM + Aprendizaje Dinámico (*)

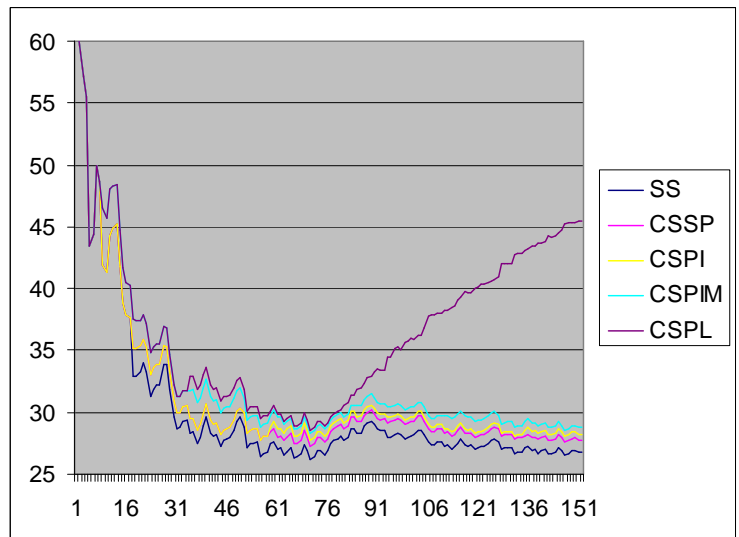


Figura 6. Ahorro de teclas combinando métodos estándar con la predicción silábica

4.- Conclusiones y Trabajos Futuros

Según investigaciones anteriores, el máximo teórico de ahorro de letras en del 75% (Sira Palazuelos, Tesis PhD), la meta es tratar de alcanzar valores superiores al 50%. En este artículo se presenta un sistema de predicción de palabras para el soporte de la escritura. Las diferentes interfaces del sistema permiten su interacción tanto con personas discapacitadas como con otras que no lo son. El sistema propuesto permite el uso de todas las aplicaciones disponibles en el mercado debido a que solamente ofrece un listado de palabras (y un teclado virtual opcional) para la escritura automática de la palabra predicha seleccionada por el usuario o el carácter seleccionado con la vista o el sistema de barrido. Además, Casandra posee otras características tales como selección automática del léxico para mejorar el rendimiento de predicción, actualización de los mismos y ajuste de pesos de los diferentes modelos. El objetivo es maximizar el rendimiento mientras que el sistema se mantiene fácil de utilizar por el usuario y sus detalles técnicos permanecen transparentes para el mismo.

Referencias

- [1] Carlberger A., Carlberger J., Magnuson T., Palazuelos S., Hunnicutt S., Aguilera S. Profet, a New Generation of Word Prediction: an Evaluation Study, Workshop on NLP for Communications Aids, ACL/EACL '97. Madrid, España, 1997.
- [2] Palazuelos S., Martin J., Arenas J., Rodino J., Aguilera S. Communication strategies using PredWin for people with disabilities. Conference and Workshop on Assistive Technology for Vision and Hearing Impaired. Castelvecchio Pascoli, Italia, 2001.
- [3] Fazly A., Hirst G.. Testing the Efficacy of Part-Of-Speech information in Word Completion. 10th EACL. Budapest, Hungary, 2003.
- [4] Gustavii E., Petterson E. A Swedish Grammar for Word Prediction. PhD thesis.
- [5] WordQ Writing Aid Software. <http://www.wordq.com/>.
- [6] Gonzalez, E., Garcia Garcia J. C. y Tulli J. C. Estudio de factibilidad de un mouse controlado por señales electrooculográficas. XV Congreso Argentino de Bioingeniería.
- [7] Chris Manning and Hinrich Schütze. Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA: May 1999.
- [8] Real Academia Española, Ortografía de la Lengua Castellana. 1826.