

ESQUEMA - RESUMEN

NICOLAS ANTONIO CAMPOS

Departamento de Lingüística Francesa

Colegio Universitario (Universidad Complutense)

Ciudad Real

Siendo consciente de mis limitaciones, ya que por una parte, mi formación lexicométrica es autodidacta y mis conocimientos en informática son pobres, y por otra, el tener que desplazarme a Madrid (Centro de Cálculo de la Universidad Complutense) para tener acceso a los medios técnicos necesarios para llevar adelante las investigaciones me obliga a ser muy modesto en cuanto a los objetivos.

Trato de realizar el estudio del vocabulario político de 200 textos (250.000 ocurrencias) de 5 publicaciones francesas de periodicidad semanal y de ámbito estatal que durante 1.968 aporten, divulguen o comenten la revolución estudiantil de mayo.

Una opción fundamental guía nuestra práctica: El texto como dato debe ser, en la medida de lo posible, fiel al documento sobre el cual trabajamos. Esto implica:

- a) Un registro exhaustivo
- b) Una precodificación reducida
- c) Una asepsia controlada

. Clasificar las formas del texto por orden alfabético indicando su frecuencia absoluta y sus referencias.

. Clasificar las mismas formas por orden de frecuencia decreciente, haciendo figurar su rango y su frecuencia relativa.

Ajustarnos rigurosamente a este punto de vista, retener todo y de la misma manera será imitar el comportamiento de un paleógrafo intentando descodificar un texto en lengua desconocida.

Salvando algunas dificultades para desambiguar el texto y procurando dar un estatuto general a los signos procederemos - de la manera siguiente:

1. Automatización del texto

- 1.1 Registro del texto
- 1.2 Registro de información complementaria al texto, fichero de parámetros.
- 1.3 Indexación
- 1.4 Delimitación y fragmentación del corpus
- 1.5 Resultado de la indexación. Corrección.

2. Estadística de frecuencias

- 2.1 Índice alfabético de formas léxicas y funcionales
- 2.2 Índice jerárquico por orden de frecuencia decreciente con indicación de F.A. y F.R.
- 2.3 Estudio de la variabilidad de frecuencia de las formas. Test apropiados.
- 2.4 Índice de reparto
- 2.5 Formas de base y formas específicas (+ y -)

3. Localización

- 3.1 Las coocurrencias y los contextos
- 3.2 Relación de las "Couples"
- 3.3 Relación de las "Couples" en contexto
- 3.4 Parejas coocurrentes

Análisis por campos, comparación de subcorpus. Análisis de las formas léxicas más frecuentes y los hapax. Ausencia/presencia.

PROCESAMIENTO DEL HABLA
Francisco Casacuberta y Enrique Vidal
CENTRO DE INFORMATICA DE LA UNIVERSIDAD DE VALENCIA

1.- INTRODUCCION.

El reconocimiento de la palabra continua es una de las facultades que caracteriza la inteligencia humana y en la actualidad se es consciente de la enorme complejidad que supone la concepción de sistemas que intenten aproximarse a las prestaciones del cerebro humano. La investigación a desarrollar en este dominio es, pues, a la vez fundamental (estudios de las propiedades y mecanismos del habla) y aplicada (desarrollo de sistemas prácticos).

En la figura adjunta se recogen esquemáticamente los proyectos en curso de desarrollo en el C.I.U.V., junto con los problemas que ellos abordan.

2.- LÍNEA FUNDAMENTAL.

En el aspecto informático de la rama fundamental, se abordan problemas de representación del conocimiento y se trabaja en el desarrollo de técnicas de Reconocimiento Sintáctico de Formas y en la aplicación de la Teoría de Conjuntos Difusos a la descripción sintáctica de las formas acústicas del habla; mientras que en el aspecto de Tratamiento de la Señal se investigan nuevos modos de caracterizar los segmentos transitorios de la señal vocal.

2.1.- RECONOCIMIENTO SINTACTICO.

En este punto se estudia, en primer lugar un sistema que permita una descripción sintáctica-difusa de los elementos fonéticos o subfonéticos del habla, esto es, partiendo de una familia indexada de conjuntos de parámetros (segmentos), extraídos mediante ventanas que se deslizan en el tiempo, se trata de obtener otra familia indexada de subconjuntos difusos de un determinado conjunto de símbolos fonéticos o subfonéticos (Etiquetado Difuso), mediante un descriptor y un traductor difuso (1).

Con el etiquetado difuso se cambia la representación paramétrica clásica por una representación simbólica de la señal vocal.

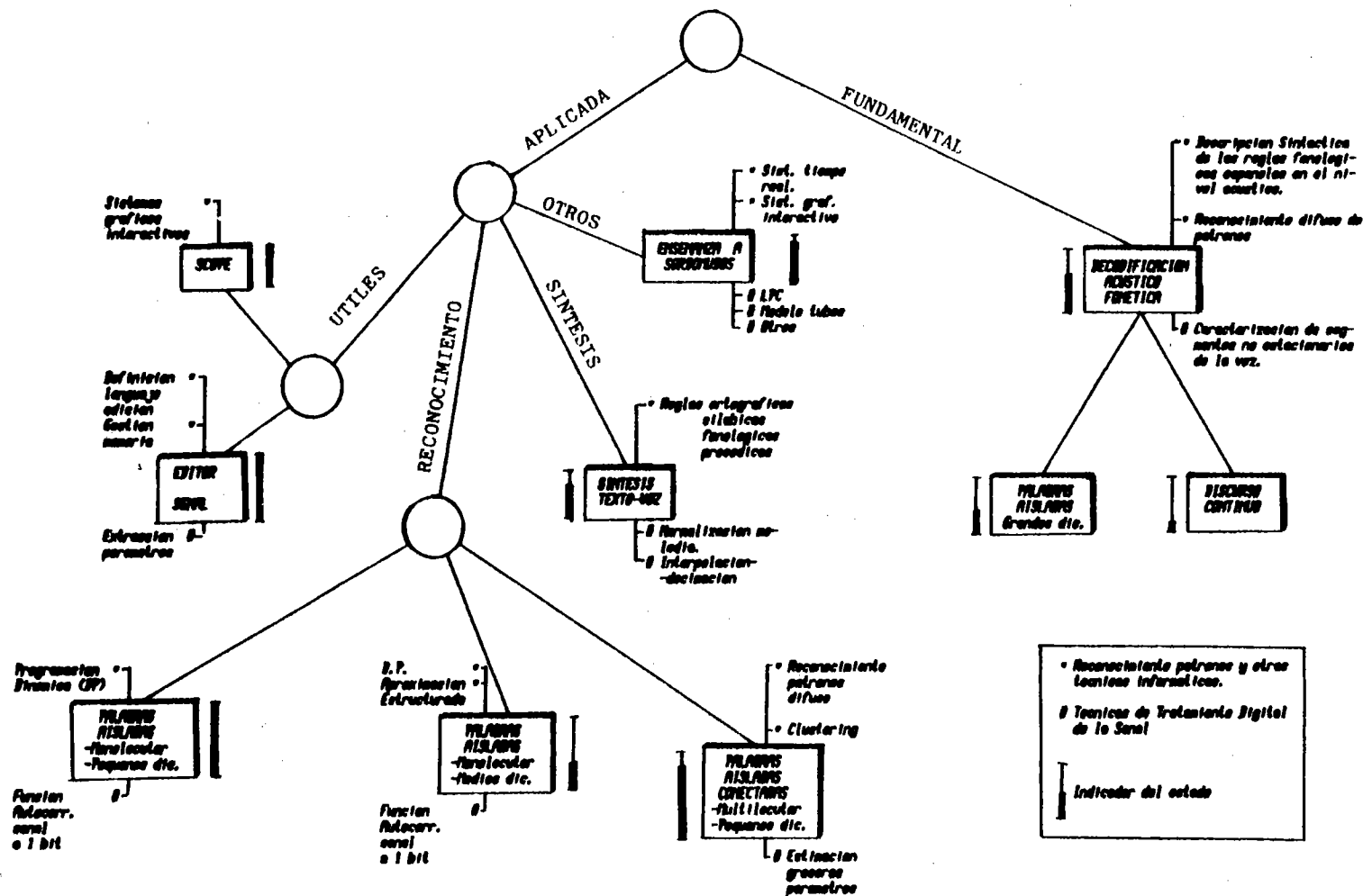
En segundo lugar se estudian los sistemas de reconocimiento propiamente dichos, esto es, partiendo de una descripción sintáctico-difusa de una palabra pronunciada y de una descripción fonológica, como representación del conocimiento, se obtiene un subconjunto difuso de palabras que forman el diccionario, donde la palabra de mayor evidencia puede ser considerada como la palabra más próxima a la pronunciada según el sistema (2).

Estos trabajos han sido aplicados a pequeños sistemas experimentales que serán descritos más adelante.

2.2.- CARACTERIZACION DE SEGMENTOS NO ESTACIONARIOS DE LA VOZ.

Uno de los problemas angulares en la decodificación acústico-fonética es el reconocimiento de los segmentos explosivos y nasales. Dichos segmentos son de gran importancia debido a su gran número en la mayoría de las lenguas, su caracterización resulta de extremada complejidad debido a su carácter no estacionario, y normalmente las soluciones que se aportan no se dan en el nivel

ESTADO DE PROYECTOS Y ACTIVIDADES DE INVESTIGACION Y RECONOCIMIENTO EN EL C.I.U.U. (1983)



acústico.

El método que se propone se basa en los estudios de la evolución en el tiempo de los parámetros de los formantes que caracterizan a las vocales siguientes al sonido consonántico (3), y en un modelo para la detección de estos parámetros en segmentos estacionarios de voz, (4), y en el cual suponemos una variación lineal de éstos; así pues, en nuestro caso, un formante, en un segmento no estacionario, viene definido por la amplitud y frecuencia central del polo y por sus correspondientes derivadas temporales. Para la obtención de dichos parámetros se han propuesto dos modelos: el primero se basa en una variación apreciable de período a período, realizándose un análisis frecuencial con dos períodos de voz (5) mientras que en el segundo se suponen fuertes variaciones dentro de un período, aplicándose un análisis frecuencial continuo en un período (6).

Aunque los resultados obtenidos son parciales, éstos hacen suponer que los modelos escogidos darán buenos resultados en un futuro próximo.

3.- LINEA APLICADA.

En esta línea se abordan proyectos prácticos en los cuales tienen aplicación inmediata los resultados de las investigaciones fundamentales. Caben destacar dos sistemas ya en funcionamiento, de Reconocimiento de Palabras Aisladas, uno monolocator y otro multilocator para diccionarios específicos. La eficacia y simplicidad de estos nuevos sistemas, permite su implementación sobre microprocesadores de propósito general, sin prácticamente ningún material adicional.

Otros proyectos abordados en esta rama abarcan temas desde la Edición Interactiva de señales vocales, hasta la Enseñanza del Habla a sordomudos, pasando por un Sistema de Síntesis sin restricciones de frases pertenecientes a cualquier texto castellano ortográficamente correcto.

3.1.- RECONOCIMIENTO DE PALABRAS AISLADAS (R.P.A.).

3.1.1.- SISTEMA DE R.P.A. MONOLOCUTOR.

Este proyecto parte de un estudio experimental realizado en nuestro Laboratorio, en el cual se demuestra que la función de autocorrelación de la señal cuantizada a 1 bit, puede ser utilizada como método de parametrización de un Sistema de Reconocimiento de Palabras Aisladas (7). Esta función es de una complejidad de cálculo pequeña y puede ser implementada con un material de bajo coste o bien en ensamblador, permitiendo obtenerla en tiempo real en un (micro) miniordenador de propósito general.

Para la fase de decisión se implementó un algoritmo de Programación Dinámica simple y rápido: el de forma simétrica sin restricciones de pendiente y con peso unidad en el camino diagonal (8).

Las tasas de reconocimiento obtenidas de este sistema son del orden de 97% para diccionarios de unas 30 palabras (9).

En este momento se está abordando el proyecto de implementar un sistema real (un mini-query por la voz), basado en el sistema aquí expuesto brevemente.

3.1.2.- SISTEMA DE R.P.A. MULTILOCUTOR BASADO EN EL ETIQUETADO DIFUSO Y EN PROGRAMACION DINAMICA.

El etiquetado difuso, estudiado en el apartado 2, utiliza, en este caso, como parámetros a la amplitud media y a la densidad de cruces por cero de un segmento de voz, y como etiquetas un conjunto de categorías fonéticas "groseras".

El tipo de representación indicado, permite, en primer lugar, la obtención del diccionario mediante traducción de las cadenas ortográficas en cadenas de símbolos que representan categorías fonéticas "groseras", utilizando las reglas fonológicas correspondientes. Y en segundo lugar, una representación contigua de aquellas cadenas que comiencen con las mismas subcadenas (9).

En la etapa de reconocimiento, propiamente dicha, se ha utilizado un algoritmo de Programación Dinámica, que aprovecha la simplicidad de dicho tipo de representación.

Las tasas de reconocimiento obtenidas con este sistema son bajas, alrededor del 70 %. Para un sistema real se pueden mejorar los resultados simplemente aumentando el poder descriptivo de la representación mediante el uso de parámetros complementarios.

3.1.3.- SISTEMA DE R.P.A. MULTILOCUTOR BASADO EN METODOS SINTACTICOS

El sistema desarrollado en este proyecto está basado en el etiquetado difuso citado en el apartado anterior, y en métodos difuso-sintácticos como sistema de decisión, utilizando una descripción sintáctica del léxico basada en las categorías acústico-fonéticas mencionadas, al cual se le asocia un algoritmo de comparación de la cadena de etiquetas difusas de entrada con la representación de las palabras del léxico. Una descripción formal de este sistema puede encontrarse en (9).

El sistema de reconocimiento descrito en este apartado ha sido implementado con una serie de modificaciones y experimentado para tres diccionarios: dígitos castellanos, dígitos catalanes y un lenguaje mini-query, y han sido obtenidas unas tasas de reconocimiento, del 85 %, resultado elevado teniendo en cuenta la "pobreza" de los parámetros utilizados.

Se han utilizado los mismos parámetros que en el sistema anterior para la obtención del etiquetado difuso, habiéndose previsto la incorporación de un tercer parámetro (densidad de cruces por cero de la señal derivada) que permitirá subdividir las categorías fonética utilizadas, y por tanto una mejor descripción del léxico.

3.2.- SINTESIS AUTOMATICA DEL CASTELLANO HABLADO.

Para contemplar el proceso inverso al Reconocimiento del Habla, se está desarrollando un proyecto de síntesis automática a partir de un texto escrito. Este proyecto se basa en decodificación de una cadena constituida por determinados elementos discretos (difonemas) concatenados y a los cuales se les ha aplicado determinadas reglas para obtener una mayor similitud al castellano hablado. El sistema utiliza un diccionario de difonemas

que consta de señales en el dominio del tiempo digitalizado (10).

El sistema se compone de una serie de etapas, algunas de las cuales se realizan en paralelo; así, a partir del texto escrito, se realiza una división silábica obteniéndose a continuación una transcripción fonética y de esta, a su vez, una transcripción difonética, con la cual se puede realizar la extracción del conjunto de difonemas del diccionario. En paralelo con las etapas descritas se obtiene a partir del texto, la información prosódica (el acento y la entonación) que una vez procesada intervienen en la concatenación de los difonemas y las variaciones de frecuencia de salida de muestras del conversor digital analógico. En la actualidad este último conjunto de etapas aun no ha sido implementado.

3.3.- AYUDA A LA ENSEÑANZA DEL HABLA A SORDOMUDOS.

Una de las aplicaciones de mayor interés social en este área de investigación, es el diseño de sistemas de ayuda a minusválidos y, en particular para el aprendizaje del habla de sordomudos.

La idea general es la de suplir la realimentación que se realiza mediante el sistema auditivo por un gráfico en el cual se muestra de alguna forma el resultado de la pronunciación de un sonido por un locutor. En nuestro caso mediante la representación en tiempo real, de un modelo de 8 tubos del aparato fonador humano, que es representado en el terminal de ordenador, teniendo a la vez la de un patrón con el mismo modelo.

El sistema desarrollado en nuestro Laboratorio está en fase de prueba con diferentes individuos sordomudos.

3.4.- UTILES

3.4.1.- SISTEMA GRAFICO INTERACTIVO: SCOPE.

SCOPE es un sistema gráfico interactivo orientado específicamente a la representación por una pantalla de rayos catódicos de señales y-z-t, y está basado en un microprocesador con un material específico para controlar la citada pantalla de bajas prestaciones. Las principales características son: flexibilidad de uso, fácil instalación en un terminal alfanumérico asíncrono y un bajo coste de implementación (11). Los dos subsistemas que componen SCOPE son: Un sistema básico y uno auxiliar. El primero, basado en el microprocesador, puede ser considerado como un dispositivo insertado en la línea asíncrona que une el terminal alfanumérico y un ordenador. El sistema auxiliar reside en el ordenador principal y permite el diálogo con el sistema base, pudiendo obtenerse copias permanentes en fichero o sobre trazador X-Y ("plotter") (11).

3.4.2.- EDITOR DE SEÑAL (EDS).

El Editor de Señal desarrollado en nuestros laboratorios es un útil de edición de señales cualesquiera, y en particular para el tratamiento de señales acústicas. Las principales aplicaciones del EDS son: creación de diccionarios de elementos acústicos y verificación de resultados en Síntesis, segmentación y marcaje manual en Reconocimiento, etc.

El EDS lleva asociado por un lado un Leguaje de Comandos que permite describir tratamientos para la manipulación de la señal, tal como inserción,

borrado, inversión, etc., así como un comando general definido por el usuario, para permitir cualquier manipulación de la señal. Los objetos del Lenguaje de Edición son representaciones de la señal acústica que admiten clasificaciones en diversos "tipos" que son referenciadas por nombres simbólicos (variables)(12).

4.- CONCLUSIONES.

Algunos de los trabajos presentados pueden darse por finalizados, sirviendo en este momento para el diseño de maquetas reales, o como herramienta de investigación. Otros en cambio están en fase de estudio y/o desarrollo, y su aplicación será a medio plazo. Todos ellos son el fruto del esfuerzo de un grupo de 6 personas que desde finales de 1980 han dedicado parte de su tiempo a la investigación e implementación en el campo del Reconocimiento y Síntesis Automática del Habla, construyendo los temas de varias tesis realizadas o en curso de elaboración.

BIBLIOGRAFIA.

- (1) VIDAL E., CASACUBERTA, F., RULOT H., SANCHIS E. " Isolated Word Recognition based on Fuzzy Labelling and Dynamic Programming" 1983 SWSSPA. Sept 83 Sitges.
- (2) VIDAL E., SANCHIS E., CASACUBERTA F. " A Speaker-Independent Isolated Word Recognition System for Specific Dictionaries" 1983 SWSSPA. Sep 83. Sitges.
- (3) RUSKE G. "Auditory Perception and its Application to Computer Analysis of Speech" Computer Analysis and Perception. Vol II. 1981.
- (4) FRIEDMAN D. "Estimation of Formant Parameters from a Sum of Pole Modeling" Proc. ICASSP-81, pp. 351-354. Apr-81.
- (5) VIDAL E., CASACUBERTA F. " Incremental Linear Model for Non Stationary Voiced Speech Segments" 11.ICA.Toulouse. Jul. 83.
- (6) CASACUBERTA F., VIDAL E. "A Continuous Linear Model for Non Stationary Analysis of Voiced Speech Transitions: an Advanced". 1983 SWSSPA. Sep. 83 Sitges.
- (7) RULOT H., VIDAL E., CASACUBERTA F. "La Función de Autocorrelación en el Reconocimiento de la Palabra" V Congreso de Inf. y Aut. Madrid 81, pp. 799-803.
- (8) RULOT H., VIDAL E., CASACUBERTA F. " Isolated Word Recognition System based the Autocorrelation Function ". 1982 PWSSPA. Oct. 82 Pavia de Varzim.
- (9) VIDAL E., CASACUBERTA F., SANCHIS E., RULOT H. " Diversas Aproximaciones al Reconocimiento de Palabras Aisladas ". V Escuela de Verano de Informática Cáceres. Jul. 83.
- (10) TORRES B., VIDAL E. " Sistema de Síntesis automática del Castellano Hablado " V Congreso de Inf. y Aut.. Madrid 81.
- (11) VIDAL E., TORRES B., CASACUBERTA F. " Un Sistema Gráfico Interactivo para la Representación de Señales Unidimensionales: SCOPE " Rev. de Informática y Automática. Pendiente de publicación.
- (12) BENEDI J. " Un Editor de Señal ". Tesina de Licenciatura. Mar. 1983. Univ. Valencia.