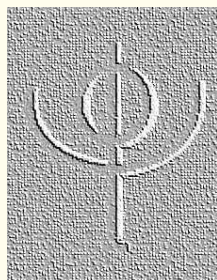


ISSN: 0325-2043



LABORATORIO DE INVESTIGACIONES SENSORIALES (LIS)

Informe XLVIII–2015

I N I G E M



CONICET

U B A

Instituto de Inmunología, Genética y Metabolismo
Córdoba 2351, Piso 9, (1121), Buenos Aires
Tel/Fax: 5950-9024
lis@fmed.uba.ar — <http://www.lis.secyt.gov.ar>

Índice

1. Introducción	1
2. Personal	1
3. Proyectos de Investigación	2
3.1. Identificación Forense de Hablantes	2
3.1.1. Trabajos terminados	2
3.2. UBACyT CM12, “Pruebas clínicas de análisis de la voz y el habla. Mejoras en la evaluación Audio-Perceptual Evaluación objetiva de la Prosodia”	2
3.2.1. Trabajos terminados	3
3.3. Desarrollo de un sistema de conversión de texto en habla para su aplicación en sistemas de telecomunicaciones	3
3.3.1. Trabajos terminados	4
3.4. Desarrollo de un sistema de reconocimiento automático del habla para el español de Argentina	7
3.4.1. Trabajos en desarrollo	7
3.5. Un rango efectivo de viscosidad	8
3.6. CONICET PIP Nro. 5897/06: Análisis de las sensaciones de dulce, agrio y amargo en soluciones puras y mezcladas en medio acuoso y alcohólico	8
3.6.1. Trabajos terminados	8
4. Docencia	9
4.1. Cursos de grado	9
5. Intercambio Científico	9
5.1. MINCYT-COLCIENCIAS	9
6. Tesis	9
6.1. Doctorales finalizadas	9
6.2. Doctorales en curso	10
7. Actividades de Divulgación	10
8. Trabajos que refieren a actividades del LIS	11
8.1. Presentaciones a congresos	11
9. Publicaciones	11
9.1. Revistas	11
9.2. Congresos	11
9.3. Informes Técnicos	12

1. Introducción

Desde su creación en el año 1968, el LIS publica un informe anual en donde se consignan las publicaciones realizadas, los trabajos en curso, la actividad docente y el intercambio científico.

Los Informes LIS están registrados bajo ISSN 0325-2043 (International Standard Serial Number), a través de Latindex¹, reconocido internacionalmente para la identificación de las publicaciones seriadas. La serie comienza con el Informe I-1968, Laboratorio de Investigaciones Sensoriales, CONICET.

En los informes aparecen siglas que referencian las sedes del LIS, primero en el Hospital Escuela (HE), luego en la Facultad de Medicina (FM) y, actualmente, en el Hospital de Clínicas (HC) de la Universidad de Buenos Aires.

Desde el año 1997, los informes también están disponibles a través del sitio web del laboratorio: <http://www.lis.secyt.gov.ar/>.

2. Personal

Investigadores

- EVIN Diego, Bioingeniero, Dr. en Ciencias de la Computación.
- GUIRAO Miguelina, Prof. Filosofía, Dra. en Psicología Experimental.
- GURLEKIAN Jorge A., Ing. Electrónico, Dr. en Medicina. Responsable del LIS.
- TORRES Humberto, BioIngeniero, Dr. en Ingeniería.

Investigadores que participan en proyectos que se desarrollan en el LIS:

- CALVIÑO Amalia M., Farmacéutica, Dra. en Bioquímica.
- GRAVANO Agustín, Licenciado y Dr. en Ciencias de la Computación.
- VACCARI María Elena, Lic. en Fonoaudiología.

Becarios doctorales

- COSSIO MERCADO Christian, Ing. en Informática, Becario CONICET. Tesista Doct UBA.
- MARTINEZ SOLER Miguel, Ing. en Informática, Becario CONICET, Tesista Doct UBA.

Becarios postdoctorales

- DE MIER Mariela Vanesa, Lic y Dra. en Letras, Becaria CONICET.

Tesistas

- SIGAL Liliana, Lic. en Fonoaudiología. Tesista de Doctorado UBA.
- TRIPODI Mónica, Lingüista. Tesista de Doctorado UBA.
- UNIVASO Pedro, Ing. Electrónico. Tesista de Doctorado UBA.

Secretaria: SOLINI Gabriela, CONICET.

¹Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal. Sitio: <http://www.latindex.unam.mx>

3. Proyectos de Investigación

3.1. Identificación Forense de Hablantes

Director: Jorge A Gurlekian, Tesistas: Miguel Martínez Soler, Pedro Univaso.

Los intercambios pasados con Gendarmería Nacional y el Ministerio Público de la Provincia de Buenos Aires permitieron comprender los requerimientos y enfoque que dan estas instituciones al tema de identificación forense. Hacia fines del 2015 la Dirección de Vinculación Tecnológica del CONICET nos invitó a participar en el comité científico para la preparación de un congreso internacional sobre Ciencia y Justicia que se desarrollará el próximo año, en virtud de la experiencia obtenida en los temas de identificación de hablantes para uso forense. La misma Dirección de Vinculación está promoviendo la formación de recursos en el área de las Facultades de Derecho de la provincia de Buenos Aires. En este período se continúa con el desarrollo de dos tesis de doctorado que emplean métodos convencionales y técnicas del estado del arte como los I-vectors para la identificación de hablantes empleando bases de datos propias y bases provistas por NIST.

3.1.1. Trabajos terminados

Data Mining applied to Forensic Speaker Identification

Univaso P., Ale J. (UCA), Gurlekian JA

Resumen:

In this paper we analyze the advantages of using data mining techniques and tools for data fusion in forensic speaker recognition. Segmental and suprasegmental features were employed in 28 different classifiers, in order to compare their performances. The selected classifiers have different learning techniques: lazy or instance-based, eager and ensemble. Two approaches were employed on the classification task: the use of all features and the use of a feature subset, selected with a gainratio methodology. The best performances, with all features, were obtained by three classifiers: Logistic Model Tree (eager), LogitBoost (ensemble) and Multi-layer Perceptron (eager). Support Vector Machine (eager) proved to be a good classifier if a Pearson VII function-based universal kernel was used. When low dimensional features were selected, ensemble classifiers exceeded the performance of all other classifiers. Segmental and tone features demonstrated the best speaker discrimination capabilities, followed by duration and quality voice features. Evaluation was performed on Argentine-Spanish voice samples from the Speech_Dat database recorded on a fixed telephone environment. Different recording sessions and channels for the test segments were added and the Z-norm procedure was applied for channel compensation.

3.2. UBACyT CM12, “Pruebas clínicas de análisis de la voz y el habla. Mejoras en la evaluación Audio-Perceptual Evaluación objetiva de la Prosodia”

Director: María E. Vaccari (UBA-Fonoaudiología), Co-director: Jorge A Gurlekian, Investigadores: Humberto Torres, Diego Evin, Liliana Sigal.

Período: 2011–2014.

3.2.1. Trabajos terminados

Comparison of two perceptual methods for evaluation of F0 perturbation

Gurlekian J.A.; Torres H.M.; Vaccari M.E.

Abstract:

Objectives. To explore perceptual evaluation of jitter produced by fundamental frequency (F0) variation in a sustained vowel /a/, using two different methods. One is based on listeners internal references and the other is based on external references provided by the experimenter. **Methods.** We used two methods: one is magnitude estimation-converging limits (ME-CL), which is close to the standard approach used by speech therapists when they use numerical estimations and their own standards, and other is intramodal matching procedure (IMP), where each matched stimulus is to be compared with a fixed-set matching stimuli. Systematic variations were introduced in vowel /a/ by Linear Prediction Coding synthesis using an F0 contour function obtained from a statistical jitter model. Six jitter values were used for each of two reference F0 values.

Three groups of listeners were tested: expert speech therapists, speech therapy students, and naive listeners. **Results.** Perceptual functions appear to be similar and linear for both methods as the theory predicts. The answers of all groups of listeners tested with ME-CL present higher standard deviations than for IMP. When subjects were tested with IMP, intrareliability and interreliability measurements show a significant improvement for both expert and naive listeners.

Conclusions. Both intraindividual and interindividual differences for expert speech therapists could be better managed when tested with an IMP than when they use numerical estimations and internal standards to evaluate vowel perturbation produced by jitter. This procedure could be the basis for the development of a clinical evaluation tool.

3.3. Desarrollo de un sistema de conversión de texto en habla para su aplicación en sistemas de telecomunicaciones

Director: Jorge Gurlekian, Co-director: Humberto Torres, Investigadores: Diego Evin, Agustín Gravano, Christian Cossio-Mercado.

Se continúa con el desarrollo iniciado con el proyecto PAE-PID 094 Se trabajó en la incorporación de mayor cantidad de datos de la locutora seleccionada. Se trabajó en la mejora de la calidad a nivel de la fusión de palabras. Los aspectos prosódicos fueron investigados en mayor detalle. Este conocimiento se volcó al conversor de texto a habla lográndose una entonación de muy alta calidad prosódica. Todas estas mejoras redundaron en un alto nivel de naturalidad como se puede apreciar en la demostración disponible en:

URL: <http://181.28.244.40:8090/LISTools/>

Se inicio la transferencia de tecnología del sistema TTS a la empresa BDT Group mediante un STAN. La empresa ha firmado en la actualidad un convenio de confidencialidad. También se presentó el sistema TTS a una segunda empresa Global News.

3.3.1. Trabajos terminados

Novel Estimation Method for Superpositional Intonation Model: From Text to Form

Torres, H.M. & Gurlekian J.A.

Abstract:

Fujisaki's intonation model parameterizes the F0 contour efficiently and giving its strong physiological basis has been successfully tested in different languages. One problem that has not been fully addressed is the extraction of the model's parameters, i.e., given a sentence, which of the model parameter values best describe its intonation. Most of the proposed methods strive to optimize the parameters so as to obtain the best fit for the F0 contour globally. In this paper we propose to use text information from the sentence as the main guide or reference for adjusting the parameters. We present a method that defines a set of rules to fix and optimize the model's parameters. Optimization never loses sight of the events of the text structure that arouse it. When text information is not enough, the algorithm predicts parameters from F0 contour and tie it to the text. The process of parameter estimation can be seen as a way to go from text information to the F0 contour. The parameter optimization is carried out to fit the F0 contour locally. Our novel approach can be implemented manually or automatically. We present examples of manual implementation and the quantitative results of the automatic one. Tested on three corpora in Spanish, English and German, our automatic method show an performance of 34 % better than other tested methods.

Estudio del Foco

Gurlekian J.A., Torres H.M., Mixdorff H. (Beuth Hochschule für Technik Berlin); Cossio-Mercado C., Güemes M.

Resumen:

Se pretende establecer una conexión entre el significado lingüístico de una emisión indicado por la modalidad y el foco con los múltiples rasgos posódicos subyacentes. De un conjunto de oraciones estructuradas estudiadas en el proyecto AMPER para el Español de Buenos Aires se analizan los datos acústicos de los parámetros tradicionales: energía, frecuencia fundamental, duración y estructura armónica y de los parámetros derivados de los comandos de acento tonal obtenidos con el modelo de entonación de Fujisaki.

Esta información acústica se correlaciona con la evaluación del foco y de las prominencias percibidas y se ensaya la detección automática de prominencia a partir de la combinación de las mediciones obtenidas.

Los resultados verifican una relación entre la máxima prominencia percibida en el objeto con las distintas condiciones de foco. Así mismo se concluye que los parámetros acústicos medidos en la sílaba como los derivados del modelo de entonación contribuyen en forma integrada a la determinación de las prominencias de los casos con un 90.78 % de certeza, utilizando el método de regresión logística. Los parámetros acústicos mejor correlacionados con la prominencia percibida son la duración de la sílaba y la amplitud del comando de acento tonal.

Acoustic Correlates of Perceived Syllable Prominence in German

Mixdorff H., Cossio-Mercado C., Hönemann A., Gurlekian J.A., Evin D., Torres H.M., Interspeech 2015, Dresden, Alemania.

URL: http://www.isca-speech.org/archive/interspeech_2015/i15.0051.html

Abstract:

This paper explores the relationship between perceived syllable prominence and the acoustic properties of a speech utterance. It is aimed at establishing a link between the linguistic meaning of an utterance in terms of sentence modality and focus and its underlying prosodic features. Applications of such knowledge can be found in computer-based pronunciation training as well as general automatic speech recognition and understanding. Our acoustic analysis confirms earlier results in that focus and sentence mode modify the fundamental frequency contour, syllabic durations and intensity. However, we could not find consistent differences between utterances produced with noncontrastive and contrastive focus, respectively. Only one third of utterances with broad focus were identified as such. Ratings of syllable prominence are strongly correlated with the amplitude of underlying accent commands, syllable duration, maximum intensity and mean harmonics-to-noise ratio.

Estimación de las funciones de costo para la selección de unidades en AROMO

Humberto Torres

Resumen:

El objetivo es estimar en forma automática las funciones de costos de selección de unidades para la voz Emilia en el sistema de conversión de texto en habla Aromo.

La síntesis por selección de unidades consiste en elegir, de entre todas las secuencias posibles de unidades que sinteticen el texto de entrada, aquella que en forma conjunta se acerque más a la secuencia predicha y la que tenga menos ruidos por la concatenación. En el sistema de conversión de texto en habla Aromo hemos implementado un algoritmo de programación dinámica que permite encontrar la mejor secuencia posible de síntesis. El método de selección de unidades se basa en evaluar cada unidad con dos métricas: el costo de concatenación y el costo objetivo. El costo de concatenación mide el ruido generado al unir dos unidades. El costo objetivo mide que tanto se parece la unidad bajo análisis a la unidad deseada. Ambos costos luego se combinan para dar una métrica de la bondad de cada unidad posible. En un informe anterior presentamos los resultados de la evaluación de distintos atributos de las unidades para la construcción de los costos. También presentamos una propuesta para de la definición de las funciones de costo. En este informe presentaremos una nueva propuesta de estimación de las funciones de costo relacionadas con los atributos de frecuencia fundamental, energía y contenido espectral. Los resultados son obtenidos corresponden a la estimación de las funciones de costo sobre la versión final del corpus Emilia.

Distribución de ocurrencias de fonemas y difonemas

Humberto Torres, Diego Evin y Jorge Gurlekian

Resumen:

Los objetivos son 1) Crear un corpus de texto de transcripciones fonéticas, 2) Analizar la distribución de los fonemas en el habla leída, 3) Analizar la distribución de los difonemas en el habla leída y 4) Estimar la cobertura del corpus Emilia.

La frecuencia con que se repiten los sonidos del habla es de interés teórico y aplicado. Desde el punto de vista teórico, la distribución de las unidades caracteriza a la lengua. Desde el punto de vista de las aplicaciones, esta información es apreciada en el campo de la percepción del habla, reconocimiento automático del habla, síntesis de habla, patologías del habla, identificación del locutor, psicoacústica, procesos cognitivos, neurolingüística, entre otros. Existen varios trabajos que reportan distribuciones de ocurrencias de fonemas para las distintas variantes del español. En particular para el español hablado en argentina, nos remitiremos al trabajo pionero de Guirao y Borzone de 1972, y de Guirao y García que realizaron el conteo

de fonemas sobre el texto de cinco libros de cuentos y novelas. Los tamaños de los corpórea utilizados hay ido evolucionando: desde 5.000 fonemas hasta 3.650.000 fonemas.

El corpus Emilia fue diseñado para la creación de una voz para el sistema Aromo de conversión de texto en habla (TTS, del inglés Text-To-Speech). En este contexto, se define la cobertura de un corpus como la capacidad de sintetizar una palabra cualquiera del idioma. Aromo realiza la síntesis del habla mediante el método de selección de unidades previamente guardadas. La unidad empleada es el difono, que se define como el segmento de habla que va desde el punto medio estable de un fono al punto medio estable del siguiente fono. Una forma de estimar si el sistema puede o no sintetizar una palabra es comparando las unidades presentes en el corpus que utiliza el sistema TTS con las unidades necesarias para realizar la tarea de síntesis.

Corpus Emilia

Resumen:

El corpus de texto Emilia fue diseñado para la tarea de conversión de texto en habla. Sus oraciones fueron creadas para contener todas las sílabas del español, en sus formas acentuadas y no acentuadas, su distribución de fonemas sigue la distribución natural de la lengua, con un número de instancias de difónos mínimo de 10. El corpus contiene 2.218 oraciones, con 29.045 palabras, de las cuales 8.919 son distintas. El corpus fue grabado por una locutora profesional para crear la versión oral. El audio fue etiquetado manualmente en varios niveles: palabras, sílabas, fonético con el sistema SAMPA (Speech Assessment Methods: Phonetic Alphabet), en clases de palabras, entre otros Error: Reference source not foundError: Reference source not foundError: Reference source not found. El corpus contiene 141.488 fonemas y 147.750 difonemas. Existen diferencias entre lo que se esperaba que emitiera la locutora y lo que finalmente pronunció. Esto se debe a tres factores: la locutora inserta pausas no marcadas que luego rompen las reglas de producción de alófonos; en algunas ocasiones la locutora cambia algunas reglas de producción de alófonos; en nombre propios de origen extranjero introduce nuevas reglas propias de transcripción de grafemas a fonemas. Esto se ve reflejado en la Error: Reference source not found, donde se puede observar las diferencias entre el etiquetado fonético teórico dado por las reglas de transcripción y lo producido por la locutora.

Corpus de texto Lana

Resumen:

El corpus de referencia Lana se construyó empleando todos los artículos publicados en la edición digital del diario La Nación de la ciudad Buenos Aires entre los años 1996 y 2006. Para obtener el texto de esos artículos se desarrolló un web crawler ad-hoc. Este programa se encarga de obtener de manera metódica las páginas web correspondientes a cada artículo, gestionando la transacción con el servidor de forma tal que el proceso de obtención de datos no perjudique el desempeño del sitio, y de filtrar el contenido textual correspondiente a cada artículo.

Como se mencionó, el propósito de este corpus es permitir un análisis estadístico a nivel grafémico e indirectamente fonético del lenguaje escrito del español de Buenos Aires. Sin embargo, las notas presentan naturalmente diversos símbolos que son propios del lenguaje escrito pero que no se corresponden a palabras. Esto implica la necesidad de llevar todos esos símbolos a sus versiones grafémicas en un proceso denominado normalización o canonización. El módulo de normalización de este material se desarrolló mediante un conjunto de scripts basados en reglas y expresiones regulares que permiten entre otras operaciones convertir símbolos numéricos a palabras, expandir abreviaturas y acrónimos, o eliminar símbolos especiales.

Por otro lado, se decidió eliminar del conjunto final de datos los artículos de la sección política del periódico. Esto se debe a que esa sección involucra muchos personajes que coyunturalmente aparecen mencionados muy frecuentemente, por ejemplo, los miembros del gobierno de turno o de la oposición, lo que para un estudio estadístico del lenguaje supone un artefacto. La versión final del corpus consiste en 55.908.107 palabras, de las cuales 288.779 son distintas.

3.4. Desarrollo de un sistema de reconocimiento automático del habla para el español de Argentina

Director: Jorge Gurlekian, Co-director: Diego Evin, Humberto Torres

3.4.1. Trabajos en desarrollo

Construcción de un corpus de datos para el reconocimiento de habla en medios de comunicación

Este trabajo se enmarca dentro del Servicio Tecnológico de Alto Nivel (STAN) “Creación de modelos acústicos y de lenguaje para el reconocimiento de habla en radio y televisión” con la empresa Global News.

Se busca crear un corpus de entrenamiento y prueba para el desarrollo de un sistema de reconocimiento automático del habla aplicado al monitoreo de radio y televisión.

Un sistema de reconocimiento automático del habla (RAH) tiene como objetivo convertir el habla presente en un segmento de audio de entrada en texto. Los sistemas actuales tienen un desempeño que dependen mucho del dominio de aplicación: en condiciones ideales de vocabulario reducido y audios con alta relación señal-ruido el nivel de reconocimiento se acerca al ideal; en condiciones menos favorables de ruido, grandes vocabularios y/o habla espontánea, el desempeño de los sistemas cae significativamente. En general, a mayor dificultad de la tarea, más complejos son los modelos utilizados y se requiere mayor cantidad de datos de entrenamiento para su construcción. En un sistema de RAH conviven dos grandes modelos: el acústico, encargado de modelar los rasgos físicos de los sonidos del habla; y las gramáticas, encargadas de caracterizar las relaciones entre las palabras.

La selección y desgravación de audios es una de las primeras tareas en la creación de un corpus para el entrenamiento y prueba de sistemas de reconocimiento de habla. Se debe transcribir literalmente lo hablado y además marcar todos los eventos acústicos. También puede ser de interés indicar el cambio de locutor, superposiciones de locutores, la presencia de ruidos, entre otros. Existen herramientas gráficas que permite leer, escuchar, ver, editar y etiquetar audios de locuciones. Este tipo de herramientas presentan varios inconvenientes de uso, por ejemplo: si los audios son de gran longitud el programa se enlentece y/o dejan de funcionar; es necesario marcar el inicio y fin de cada turno de habla antes de desgrabar; es engorroso moverse entre los distintos campos de etiquetado; entre otros problemas detectados. Una alternativa es segmentar los audios antes de su desgrabación manual. El criterio de segmentación puede tener como motivación: pausas, excesivo ruido, por tiempo, por eventos acústicos, entre otros. Una vez realizada la segmentación, los audios resultantes pueden ser guardados en archivos independientes, facilitando su manipulación y desgrabación.

En este trabajo nos propusimos crear un corpus para entrenar y evaluar un sistema automático de reconocimiento de habla en audios de radio y televisión. Para ello se crearon dos herramientas: una que permite segmentar los archivos de audio, y otra que permite su desgrabación

y etiquetado. La creación del corpus se encuentra en proceso.

3.5. Un rango efectivo de viscosidad

En este trabajo se intenta determinar el rango de viscosidad física que se correlaciona en forma efectiva y sin ambigüedades con la viscosidad percibida. La búsqueda de un rango efectivo conduce a dos interrogantes. Si los valores bajos de la escala física se correlacionan más bien con la fluidez y los altos más bien con la consistencia (entendida como la cualidad de la materia que resiste sin romperse ni deformarse fácilmente) que con la viscosidad. Se utilizarán catorce muestras de siliconas (silicone oils newtonian fluids from Brookfield Eng. Lab., USA) en un rango de valores de 2 a 1 000 000 centipoises (cps). Los panelistas recibirán las muestras en tubos y las evaluarán observando el movimiento de las soluciones. En una sección preliminar los panelistas darán juicios cualitativos. Cada uno dividirá las muestras en grupos de acuerdo a las cualidades (gradientes de viscosidad) que percibe. En las secciones siguientes darán juicios cuantitativos. Se realizarán experimentos en los que se aplicarán diferentes métodos psicofísicos y se observará si las funciones perceptuales que se obtengan se diferencian de acuerdo a los subrangos de la viscosidad física.

3.6. CONICET PIP Nro. 5897/06: Análisis de las sensaciones de dulce, agrio y amargo en soluciones puras y mezcladas en medio acuoso y alcohólico

Dirección: Miguelina Guirao, Codirección: Amalia Mirta Calviño

3.6.1. Trabajos terminados

Bitter taste modifications by mixing caffeine with ethanol

Guirao M., Calviño A., and Evin D.

Abstract:

Background: The majority of psychophysical research on bitter taste ethanol mixture have focused on complex models or on a single compound like quinine. Since different bitter compounds give different perceptual responses previous findings may not generalize to ethanol caffeine mixture. Important as it is no much data on the interaction between these two chemicals is available.

Objective: This study examine the potential changes in intensity and duration of aftertaste produced in the taste of caffeine when mixed with ethanol.

Methods: The psychophysical methods of Pair Comparison and Magnitude Estimation were applied to quantify the bitterness intensity of seven concentrations of caffeine (3, 6, 12, 25, 35, 55 and 80 mM) in water solutions tasted alone and mixed with two (8% and 15%) ethanol levels in water. The Time-Intensity Method was used to asses intensity and aftertaste of two caffeine (6 and 55 mM) solutions unmixed and mixed with the same ethanol levels.

Results: When taste intensity was rated both levels of ethanol increased bitterness at the lower half part of the caffeine range. In the upper part (up to 35 mM) the bitter taste intensity remained unchanged. When assessing intensity/time responses the mixture gave longer aftertaste than caffeine alone. The difference between the effect of the two ethanol levels is clear in the weak (6 mM) but less pronounced in the stronger (55 mM) caffeine solution.

Conclusions: The addition of ethanol to caffeine enhanced the duration of bitter aftertaste.

Taste intensity was increased at weak and moderate concentrations but higher concentrations were unaffected.

4. Docencia

4.1. Cursos de grado

Dr. Humberto M. Torres

Profesor Adjunto de la cátedra *Señales y Sistemas*, Facultad de Ingeniería UBA. Desde el 26 de Septiembre de 2011.

Dr. Jorge A Gurlekian

Dictado del Seminario *Laboratorio de Voz*. Facultad de Medicina. Área Fonoaudiología. Desde el 2001.

Ing. Christian Cossio Mercado

Jefe de Trabajos Prácticos (JTP) en las materias *Paradigmas de Lenguajes de Programación y Teoría de Lenguajes*, Departamento de Computación, Facultad de Ciencias Exactas y Naturales, UBA. Desde marzo de 2015.

5. Intercambio Científico

5.1. MINCYT-COLCIENCIAS

Proyecto Bilateral Cooperación Científico-Tecnológica entre el Consejo Nacional de Investigaciones Científicas y Técnicas, el Ministerio de Ciencia, Tecnología e Innovación Productiva de la República Argentina (MINCYT) y el Departamento Administrativo de Ciencia, Tecnología e Innovación de Colombia (COLCIENCIAS) Id. CO/14/02. “Determinación del Riesgo Vocal: verificación de los pesos relativos de los componentes subjetivos y objetivos”. Período: 2015-2016.

6. Tesis

6.1. Doctorales finalizadas

Diagnostico diferencial de pacientes con movimientos anormales laringeos, complementacion entre el diagnostico neurológico y los resultados que brinda el abordaje otorrino-fonoaudiologico

Tesista: Liliana Sigal

Universidad de Buenos Aires, Facultad de Medicina

Director : Dr. Jorge A. Gurlekian.

Consejero : Profesor Dr Federico Micheli

Calificación: Sobresaliente

6.2. Doctorales en curso

Evaluación Automática de Calidad del Habla Artificial

Tesista: Christian Cossio Mercado

Directores: Dr. José Castaño (FCEyN-UBA) y Dr. Jorge Gurlekian

Consejero de Estudios: Dr. Agustín Gravano (FCEyN-UBA)

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales

Reconocimiento forense de hablantes mediante el uso de información de alto nivel y metadatos

Tesista: Miguel Martínez Soler

Directores: Dr. Jorge A. Gurlekian (CONICET) y Dr. Agustín Gravano (FCEyN-UBA)

Consejero de Estudios: Dr. Diego Garbervetsky (FCEyN-UBA)

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales

Reconocimiento automático de hablantes empleando información de largo plazo

Tesista: Pedro Univaso

Director: Dr. Jorge A. Gurlekian

Universidad de Buenos Aires, Facultad de Ingeniería.

7. Actividades de Divulgación

- Gurlekian Jorge. Entrevista en Noticias del sitio de CONICET. 21 de julio de 2015. Disponible en <http://www.conicet.gov.ar/te-estoy-hablando-a-vos-2/>.
- Cossio-Mercado, Christian. “Todo con afecto (¡y también la Computación!)”, Semana de la Computación 2015, Facultad de Ciencias Exactas y Naturales, UBA. Charla sobre la Computación Afectiva (*Affective Computing*), y su utilización para poder conocer el estado interno de los seres humanos. 18 de Junio de 2015.
- Evin Diego. Entrevista radial, para el programa “Se va la mañana” de Radio Güemes (AM 1050), conducido por Sebastián Kohen, para hablar sobre “El desarrollo de sistemas de reconocimiento del habla en el Laboratorio de Investigaciones Sensoriales”. 29 de julio de 2015.
- Evin Diego. Entrevista radial vía telefónica para el programa “Mañana Regional”, conducido por Luisa Gimenez de Radio San Genaro (LT 23 AM1550), Santa Fe, Argentina, sobre el tema “Un sistema de reconocimiento de habla que optimizará la relación con la tecnología”. 12 de agosto de 2015.
- Evin Diego. Entrevista radial vía telefónica para el programa “Te Estoy Hablando A Vos”, conducida por Guillermo Arsuaga de Radio Saladillo (FM 90.7), Provincia de Buenos Aires, Argentina, sobre el tema “Tecnología y Reconocimiento del Habla”. 27 de julio de 2015.
- Evin Diego. Entrevista televisiva para el programa “Vivo en Argentina”, conducido por Sergio Goycochea y Maby Wells, en la TV Pública de Argentina, sobre el tema

“Inteligencia Artificial y Lenguaje”. 3 de agosto de 2015.

Disponible en <https://www.youtube.com/watch?v=8orvBmqYPS8>.

8. Trabajos que refieren a actividades del LIS

8.1. Presentaciones a congresos

- Piñeda, M. A., “Miguelina Guirao y Claribel Barbenza: negociaciones entre la cultura humanista y la cultura científica de la psicología argentina de la segunda mitad del siglo XX”. XVI Encuentro Argentino de Historia de la Psiquiatría, la Psicología y el Psicoanálisis, Mar del Plata, 2015.

9. Publicaciones

9.1. Revistas

- Univaso P., Ale J., Gurlekian J.A., “Data Mining applied to Forensic Speaker Identification”. *IEEE Latin America Transactions*, vol 13:4, 2015.
- Torres H.M., Gurlekian J.A., “Novel Estimation Method for Superpositional Intonation Model: From Text to Form”. *IEEE Transactions on Audio Speech and Language Processing*. Año 2015, mes 11. En prensa.
- Guirao M., Evin D., “Efectos del Etanol en el Gusto de la Sacarosa y del Ácido Cítrico”. *La Alimentación Latinoamericana*, vol. 321, pp.56–61, Publitec, Buenos Aires, 2015.
- Gurlekian J., Torres H., and Vaccari M., “Comparison of two methods for perceptual evaluation of F0 perturbation”. En prensa en *Journal of Voice*. 2015.

9.2. Congresos

- Alvarez, A., Evin, D., Verrastro, S., “Implementation of a Speech Recognition System in a DSC”. En *Memorias de la XVI Reunión de Trabajo en Procesamiento de la Información y Control*, Córdoba, 2015.
- Mixdorff H., Cossio-Mercado C., Hönemann A., Gurlekian J., Evin D., Torres H., “Acoustic Correlates of Perceived Syllable Prominence in German”. *Proceedings of the 16th Annual Conference of the International Speech Communication Association*, Dresden 2015, pp. 51–55.
Disponible en http://www.isca-speech.org/archive/interspeech_2015/papers/i15_0051.pdf.
- Hadad A., Solano A., Evin D., Drozdowicz B., “Temporal Comparison of Arrhythmia States Sequence”. *VI Latin American Congress on Biomedical Engineering, CLAIB 2014*, vol. 49, pp. 773–776, Springer International, Suiza, 2015.
Disponible en http://link.springer.com/chapter/10.1007/978-3-319-13117-7_197.
- Hadad A., Evin D., Drozdowicz B., “Preprocessing Unbalanced Data Set Based on Self-organizing Neural Networks”. *VI Latin American Congress on Biomedical Engineering, CLAIB 2014*, vol. 49, pp. 777–780, Springer International, Suiza, 2015.
Disponible en http://link.springer.com/chapter/10.1007/978-3-319-13117-7_198.

- Evin D., Hadad A., Solano A., Drozdowicz B., “Segmentation Fusion Techniques with Application to Plenoptic Images: A Survey”. *Proceedings del XX Congreso Argentino de Bioingeniería y IX Jornadas de Ingeniería Clínica*, San Nicolás, 2015. Disponible en <http://www.sabi2015.gadib.com.ar/actas/actas.pdf>.

9.3. Informes Técnicos

- Torres H., Evin D. y Gurlekian J., “Distribución de ocurrencias de fonemas y difonemas”. Laboratorio de Investigaciones Sensoriales, LIS, INIGEM, CONICET-UBA. Julio de 2015.
- Gurlekian J. y Torres H., “Sistema de Evaluación Audio-Perceptual EVAPER”. Laboratorio de Investigaciones Sensoriales, LIS, INIGEM, CONICET-UBA. Agosto de 2015.
- Gurlekian J., Evin D. y Torres H., “Nuevas consignas para la desgrabación y etiquetado”. Laboratorio de Investigaciones Sensoriales, LIS, INIGEM, CONICET-UBA. Septiembre de 2015.
- Evin D.A., Torres H.M., Gurlekian J.A., “Generación de Modelos de Entrenamiento para un Sistema de Reconocimiento de Habla”. Laboratorio de Investigaciones Sensoriales, LIS, INIGEM, CONICET-UBA. Destinatario: Global News. Octubre de 2015.
- Torres H., Evin D. y Gurlekian J., “Estado de avance de la construcción de un corpus de datos para el reconocimiento de habla en medios de comunicación”. Laboratorio de Investigaciones Sensoriales, LIS, INIGEM, CONICET-UBA. Destinatario: Global News. Noviembre de 2015.
- Torres H., “Estimación de las funciones de costo para la selección de unidades en AROMO”. Laboratorio de Investigaciones Sensoriales, LIS, INIGEM, CONICET-UBA. Noviembre de 2015.