

# Handbook of Research on Natural Language Processing and Smart Service Systems

Rodolfo Abraham Pazos–Rangel  
*Tecnológico Nacional de México, Mexico & Instituto Tecnológico de Ciudad  
Madero, Mexico*

Rogelio Florencia–Juarez  
*Universidad Autónoma de Ciudad Juárez, Mexico*

Mario Andrés Paredes–Valverde  
*Tecnológico Nacional de México, Mexico & Instituto Tecnológico de Orizaba,  
Mexico*

Gilberto Rivera  
*Universidad Autónoma de Ciudad Juárez, Mexico*



A volume in the Advances in Computational  
Intelligence and Robotics (ACIR) Book Series

Published in the United States of America by  
IGI Global  
Engineering Science Reference (an imprint of IGI Global)  
701 E. Chocolate Avenue  
Hershey PA, USA 17033  
Tel: 717-533-8845  
Fax: 717-533-8661  
E-mail: [cust@igi-global.com](mailto:cust@igi-global.com)  
Web site: <http://www.igi-global.com>

Copyright © 2021 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Names: Pazos-Rangel, Rodolfo Abraham, 1951- editor.  
Title: Handbook of research on natural language processing and smart service systems / Rodolfo Abraham Pazos-Rangel, Rogelio Florencia-Juarez, Mario Andrés Paredes-Valverde, Gilberto Rivera, editors.

Description: Hershey, PA : Engineering Science Reference, an imprint of IGI Global, [2020] | Includes bibliographical references and index. | Summary: "This book is a collection of innovative research on the integration and development of intelligent software tools and their various applications within professional environments"-- Provided by publisher.

Identifiers: LCCN 2019058351 (print) | LCCN 2019058352 (ebook) | ISBN 9781799847304 (hardcover) | ISBN 9781799847311 (ebook)

Subjects: LCSH: Natural language processing (Computer science) | Natural language generation (Computer science) | Computational linguistics.

Classification: LCC QA76.9.N38 H3645 2020 (print) | LCC QA76.9.N38 (ebook) | DDC 006.3/5--dc23

LC record available at <https://lcn.loc.gov/2019058351>

LC ebook record available at <https://lcn.loc.gov/2019058352>

This book is published in the IGI Global book series Advances in Computational Intelligence and Robotics (ACIR) (ISSN: 2327-0411; eISSN: 2327-042X)

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: [eresources@igi-global.com](mailto:eresources@igi-global.com).

## List of Contributors

<b>Aguirre L., Marco A.</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	289
<b>Aldana-Bobadilla, Edwin</b> / <i>Conacyt, Mexico &amp; Cinvestav Tamaulipas, Mexico</i> .....	393
<b>Almanza Ortega, Nelva Nely</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Tlalnepantla, Mexico</i> .....	289
<b>Alor-Hernández, Giner</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	135
<b>Ameer, Iqra</b> / <i>Instituto Politécnico Nacional, Mexico</i> .....	245
<b>Bonilla, Juan Carlos</b> / <i>Universidad Autónoma del Estado de Morelos, Mexico</i> .....	266
<b>Bustos-López, Maritza</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	445
<b>C., Namrata Mahender</b> / <i>Dr. Babasaheb Ambedkar Marathwada University, India</i> .....	46
<b>Castillo, German</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	196
<b>Castro-Pérez, Karina</b> / <i>Tecnológico Nacional de México, Mexico &amp; IT Orizaba, Mexico</i> .....	445
<b>Contreras-Masse, Roberto</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Juárez, Mexico</i> .....	180, 266
<b>Fernández-Avelino, Jesús</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	135
<b>Florencia-Juárez, Rogelio</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	1
<b>Frausto Solís, Juan</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	70
<b>García, Alonso</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	309
<b>García, Vicente</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	427, 481
<b>Gaspar, Juana</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	1
<b>Gelbukh, Alexander</b> / <i>Instituto Politécnico Nacional, Mexico</i> .....	157
<b>González, Martha Victoria</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	101, 309
<b>González-Barbosa, Juan Javier</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	70, 196
<b>Guzmán Mendoza, José Eder</b> / <i>Universidad Politécnica de Aguascalientes, Mexico</i> .....	327
<b>Hernández Gómez, Antonio</b> / <i>Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i> ..	289
<b>Jiménez, Rafael</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	427, 481
<b>Kumari, Namrata</b> / <i>National Institute of Technology, Hamirpur, India</i> .....	368
<b>Lopez Contreras, Irvin Raul</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	379

<b>López, Abraham</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	227, 427
<b>Lopez-Arevalo, Ivan</b> / <i>CINVESTAV Tamaulipas, Mexico</i> .....	393
<b>López-Orozco, Francisco</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	31, 309
<b>Lopez-Veyna, Jaime I.</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Zacatecas, Mexico</i> .....	393
<b>Mar, Ricardo</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	347
<b>Martínez F., José A.</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	1, 157, 196
<b>Martinez, Marcos E.</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	31
<b>Martinez-Rodriguez, Jose L.</b> / <i>Universidad Autónoma de Tamaulipas, Mexico</i> .....	393
<b>Mejia, Jose</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	180, 266
<b>Mendoza Carreón, Alejandra</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	379, 427
<b>Montes Rivera, Martín</b> / <i>Universidad Politécnica de Aguascalientes, Mexico</i> .....	327
<b>Ochoa, Alberto</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	157, 180, 266, 327
<b>Oliva, Diego</b> / <i>Universidad de Guadalajara, Mexico</i> .....	180
<b>Olmos-Sánchez, Karla</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	31, 481
<b>Ortiz Hernandez, Javier</b> / <i>Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i> .....	289
<b>Paredes-Valverde, Mario Andrés</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	135
<b>Pazos-Rangel, Rodolfo A.</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	1
<b>Pérez Ortega, Joaquín</b> / <i>Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i> .....	289
<b>Ponce Gallegos, Julio César</b> / <i>Universidad Autónoma de Aguascalientes, Mexico</i> .....	327
<b>Ponce, Alan</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	427, 481
<b>Porras, Raul</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	227
<b>Porras, Raúl</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	347
<b>Ramirez López, Carlos Manuel</b> / <i>Universidad Politécnica de Aguascalientes, Mexico</i> .....	327
<b>Requejo Flores, Alejandro</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	227, 347
<b>Rios-Alvarado, Ana B.</b> / <i>Universidad Autónoma de Tamaulipas, Mexico</i> .....	393
<b>Rivera, Gilberto</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	1
<b>Rodas-Osollo, Jorge</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	379, 481
<b>Rodríguez-Mazahua, Lisbeth</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	445
<b>Ruiz, Alejandro</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	227, 347
<b>Salas-Zárate, María del Pilar</b> / <i>Tecnológico Nacional de México, Mexico &amp; ITS Teziutlán, Mexico</i> .....	445
<b>Sánchez-Cervantes, José Luis</b> / <i>CONACYT, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	445
<b>Sánchez-Hernández, Juan Paulo</b> / <i>Universidad Politécnica del Estado de Morelos, Mexico</i> .....	70
<b>Sánchez-Morales, Laura Nely</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Orizaba, Mexico</i> .....	135
<b>Sanchez-Solís, Julia Patricia</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	70
<b>Sánchez-Solís, Julia Patricia</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	31
<b>Sayyed, Sanah Nashir</b> / <i>Dr. Babasaheb Ambedkar Marathwada University, India</i> .....	46
<b>Sidorov, Grigori</b> / <i>Instituto Politécnico Nacional, Mexico</i> .....	245
<b>Singh, Pardeep</b> / <i>National Institute of Technology, Hamirpur, India</i> .....	368
<b>Varela, Maritza</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	101

<b>Varela, Martiza Concepción</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	379
<b>Vega Villalobos, Andrea</b> / <i>Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i> .....	289
<b>Verastegui, Andres</b> / <i>Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i> .....	157
<b>Villanueva-Mendoza, Ossiel</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	101
<b>Zamora, Lucero</b> / <i>Universidad Autónoma de Ciudad Juárez, Mexico</i> .....	101, 309
<b>Zavala Díaz, Crispín</b> / <i>Universidad Autónoma del Estado de Morelos, Mexico</i> .....	289

# Table of Contents

<b>Foreword</b> .....	xxiv
<b>Preface</b> .....	xxv
<b>Acknowledgment</b> .....	xxx

## **Section 1** **Smart Interactive Systems**

### **Chapter 1**

Natural Language Interfaces to Databases: A Survey on Recent Advances .....	1
<i>Rodolfo A. Pazos-Rangel, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	
<i>Gilberto Rivera, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>José A. Martínez F., Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	
<i>Juana Gaspar, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	
<i>Rogelio Florencia-Juárez, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

### **Chapter 2**

Mispronunciation Detection and Diagnosis Through a Chatbot .....	31
<i>Marcos E. Martinez, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Francisco López-Orozco, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Karla Olmos-Sánchez, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Julia Patricia Sánchez-Solís, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

### **Chapter 3**

Story Summarization Using a Question-Answering Approach .....	46
<i>Sanah Nashir Sayyed, Dr. Babasaheb Ambedkar Marathwada University, India</i>	
<i>Namrata Mahender C., Dr. Babasaheb Ambedkar Marathwada University, India</i>	

#### **Chapter 4**

Two New Challenging Resources to Evaluate Natural Language Interfaces to Databases  
Generated Based on Geobase and Geoquery ..... 70

*Juan Javier González-Barbosa, Tecnológico Nacional de México, Mexico & Instituto  
Tecnológico de Ciudad Madero, Mexico*

*Juan Frausto Solís, Tecnológico Nacional de México, Mexico & Instituto Tecnológico de  
Ciudad Madero, Mexico*

*Juan Paulo Sánchez-Hernández, Universidad Politécnica del Estado de Morelos, Mexico*

*Julia Patricia Sanchez-Solís, Universidad Autónoma de Ciudad Juárez, Mexico*

#### **Chapter 5**

Chatbot for the Improvement of Conversational Skills of Japanese Language Learners ..... 101

*Ossiel Villanueva-Mendoza, Universidad Autónoma de Ciudad Juárez, Mexico*

*Martha Victoria González, Universidad Autónoma de Ciudad Juárez, Mexico*

*Maritza Varela, Universidad Autónoma de Ciudad Juárez, Mexico*

*Lucero Zamora, Universidad Autónoma de Ciudad Juárez, Mexico*

#### **Chapter 6**

Developing Chatbots for Supporting Health Self-Management ..... 135

*Jesús Fernández-Avelino, Tecnológico Nacional de México, Mexico & Instituto Tecnológico  
de Orizaba, Mexico*

*Giner Alor-Hernández, Tecnológico Nacional de México, Mexico & Instituto Tecnológico de  
Orizaba, Mexico*

*Mario Andrés Paredes-Valverde, Tecnológico Nacional de México, Mexico & Instituto  
Tecnológico de Orizaba, Mexico*

*Laura Nely Sánchez-Morales, Tecnológico Nacional de México, Mexico & Instituto  
Tecnológico de Orizaba, Mexico*

#### **Chapter 7**

Issues in the Syntactic Parsing of Queries for a Natural Language Interface to Databases ..... 157

*Alexander Gelbukh, Instituto Politécnico Nacional, Mexico*

*José A. Martínez F., Tecnológico Nacional de México, Mexico & Instituto Tecnológico de  
Ciudad Madero, Mexico*

*Andres Verastegui, Tecnológico Nacional de México, Mexico & Instituto Tecnológico de  
Ciudad Madero, Mexico*

*Alberto Ochoa, Universidad Autónoma de Ciudad Juárez, Mexico*

#### **Chapter 8**

Preservation of Cultural Heritage in an Ethnic Minority Using Internet of Things and Smart  
Karaoke ..... 180

*Alberto Ochoa, Universidad Autónoma de Ciudad Juárez, Mexico*

*Roberto Contreras-Masse, Tecnológico Nacional de México, Mexico & Instituto Tecnológico  
de Ciudad Juárez, Mexico*

*Jose Mejia, Universidad Autónoma de Ciudad Juárez, Mexico*

*Diego Oliva, Universidad de Guadalajara, Mexico*

## Chapter 9

Interface for Composing Queries That Include Subqueries for Complex Databases .....	196
<i>José A. Martínez F., Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	
<i>Juan Javier González-Barbosa, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	
<i>German Castillo, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	

## Section 2 Text Analytics Systems

## Chapter 10

News Classification to Notify About Traffic Incidents in a Mexican City.....	227
<i>Alejandro Requejo Flores, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Alejandro Ruiz, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Abraham López, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Raul Porras, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

## Chapter 11

Author Profiling Using Texts in Social Networks.....	245
<i>Iqra Ameer, Instituto Politécnico Nacional, Mexico</i>	
<i>Grigori Sidorov, Instituto Politécnico Nacional, Mexico</i>	

## Chapter 12

A Comparison of Personality Prediction Classifiers for Personnel Selection in Organizations Based on Industry 4.0 .....	266
<i>Roberto Contreras-Masse, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Juárez, Mexico</i>	
<i>Juan Carlos Bonilla, Universidad Autónoma del Estado de Morelos, Mexico</i>	
<i>Jose Mejia, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Alberto Ochoa, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

## Chapter 13

Improving the K-Means Clustering Algorithm Oriented to Big Data Environments.....	289
<i>Joaquín Pérez Ortega, Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i>	
<i>Nelva Nely Almanza Ortega, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Tlalnepantla, Mexico</i>	
<i>Andrea Vega Villalobos, Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i>	
<i>Marco A. Aguirre L., Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Ciudad Madero, Mexico</i>	
<i>Crispín Zavala Díaz, Universidad Autónoma del Estado de Morelos, Mexico</i>	
<i>Javier Ortiz Hernandez, Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i>	
<i>Antonio Hernández Gómez, Tecnológico Nacional de México, Mexico &amp; CENIDET, Mexico</i>	



## Chapter 14

Pronominal Anaphora Resolution on Spanish Text .....	309
<i>Alonso García, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Martha Victoria González, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Francisco López-Orozco, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Lucero Zamora, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

## Chapter 15

Geospatial Situation Analysis for the Prediction of Possible Cases of Suicide Using EBK: A Case Study in the Mexican State of Aguascalientes .....	327
<i>Carlos Manuel Ramírez López, Universidad Politécnica de Aguascalientes, Mexico</i>	
<i>Martín Montes Rivera, Universidad Politécnica de Aguascalientes, Mexico</i>	
<i>Alberto Ochoa, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Julio César Ponce Gallegos, Universidad Autónoma de Aguascalientes, Mexico</i>	
<i>José Eder Guzmán Mendoza, Universidad Politécnica de Aguascalientes, Mexico</i>	

## Chapter 16

Location Extraction to Inform a Spanish-Speaking Community About Traffic Incidents.....	347
<i>Alejandro Requejo Flores, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Alejandro Ruiz, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Ricardo Mar, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Raúl Porras, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

## Chapter 17

Text Summarization and Its Types: A Literature Review .....	368
<i>Namrata Kumari, National Institute of Technology, Hamirpur, India</i>	
<i>Pardeep Singh, National Institute of Technology, Hamirpur, India</i>	

## Chapter 18

Extractive Text Summarization Methods in the Spanish Language .....	379
<i>Irvin Raul Lopez Contreras, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Alejandra Mendoza Carreón, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Jorge Rodas-Osollo, Universidad Autónoma de Ciudad Juárez, Mexico</i>	
<i>Martiza Concepción Varela, Universidad Autónoma de Ciudad Juárez, Mexico</i>	

### Section 3 Text Mining Systems

## Chapter 19

NLP and the Representation of Data on the Semantic Web .....	393
<i>Jose L. Martinez-Rodriguez, Universidad Autónoma de Tamaulipas, Mexico</i>	
<i>Ivan Lopez-Arevalo, CINVESTAV Tamaulipas, Mexico</i>	
<i>Jaime I. Lopez-Veyna, Tecnológico Nacional de México, Mexico &amp; Instituto Tecnológico de Zacatecas, Mexico</i>	
<i>Ana B. Rios-Alvarado, Universidad Autónoma de Tamaulipas, Mexico</i>	
<i>Edwin Aldana-Bobadilla, Conacyt, Mexico &amp; Cinvestav Tamaulipas, Mexico</i>	

**Chapter 20**

Opinion Mining for Instructor Evaluations at the Autonomous University of Ciudad Juarez..... 427  
*Rafael Jiménez, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Vicente García, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Abraham López, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Alejandra Mendoza Carreón, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Alan Ponce, Universidad Autónoma de Ciudad Juárez, Mexico*

**Chapter 21**

An Opinion Mining Approach for Drug Reviews in Spanish..... 445  
*Karina Castro-Pérez, Tecnológico Nacional de México, Mexico & IT Orizaba, Mexico*  
*José Luis Sánchez-Cervantes, CONACYT, Mexico & Instituto Tecnológico de Orizaba, Mexico*  
*María del Pilar Salas-Zárate, Tecnológico Nacional de México, Mexico & ITS Teziutlán, Mexico*  
*Maritza Bustos-López, Tecnológico Nacional de México, Mexico & Instituto Tecnológico de Orizaba, Mexico*  
*Lisbeth Rodríguez-Mazahua, Tecnológico Nacional de México, Mexico & Instituto Tecnológico de Orizaba, Mexico*

**Chapter 22**

Identifying Suggestions in Airline-User Tweets Using Natural Language Processing and Machine Learning ..... 481  
*Rafael Jiménez, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Vicente García, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Karla Olmos-Sánchez, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Alan Ponce, Universidad Autónoma de Ciudad Juárez, Mexico*  
*Jorge Rodas-Osollo, Universidad Autónoma de Ciudad Juárez, Mexico*

**Compilation of References** ..... 499

**About the Contributors** ..... 542

**Index**..... 552

## Chapter 2

# Mispronunciation Detection and Diagnosis Through a Chatbot

**Marcos E. Martinez**

 <https://orcid.org/0000-0002-9777-6395>

*Universidad Autónoma de Ciudad Juárez, Mexico*

**Francisco López-Orozco**

*Universidad Autónoma de Ciudad Juárez, Mexico*

**Karla Olmos-Sánchez**

 <https://orcid.org/0000-0002-9145-6761>

*Universidad Autónoma de Ciudad Juárez, Mexico*

**Julia Patricia Sánchez-Solís**

*Universidad Autónoma de Ciudad Juárez, Mexico*

## ABSTRACT

*The interaction between humans and machines has evolved; thus, the idea of being able to communicate with computers as we usually do with other people is becoming increasingly closer to coming true. Nowadays, it is common to come across intelligent systems named chatbots, which allow people to communicate by using natural language to hold conversations related to a specific domain. Chatbots have gained popularity in different kinds of sectors, such as customer service, marketing, sales, e-commerce, e-learning, travel, and even in education itself. This chapter aims to present a chatbot-based approach to learning English as a second language by using computer-assisted language learning systems.*

## INTRODUCTION

During day-to-day activities, human beings make use of natural language. Something that characterizes natural language is its ambiguity, especially when it is expressed in written format. Hence, Artificial Intelligence (AI) community has been extensively researched and developed techniques, algorithms, and tools in order to improve the human-computer interaction. Natural Language Processing (NLP) arises in

DOI: 10.4018/978-1-7998-4730-4.ch002

### **Mispronunciation Detection and Diagnosis Through a Chatbot**

is possible to replace the DNN-HMM ASR model and GOP algorithm for a CTC-ASR model to do mispronunciation detection and diagnosis in order to have a more fluent conversation with the chatbot.

## **REFERENCES**

- Abdul-Kader, S., & Woods, J. (2015). Survey on Chatbot Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications*, 6(7), 72–80. doi:10.14569/ijacsa.2015.060712
- Banerjee, A., Dubey, A., Menon, A., Nanda, S., & Chand Nandi, G. (2018). *Speaker Recognition using Deep Belief Networks*. Retrieved from <https://arxiv.org/ftp/arxiv/papers/1805/1805.08865.pdf>
- Buschmeier, H., & Wlodarczak, M. (2013). TextGridTools: A TextGrid Processing and Analysis Toolkit for Python. In *Tagungsband Der 24. Konferenz Zur Elektronischen Sprachsignalverarbeitung (ESSV 2013)* (pp.152–157). Academic Press.
- Chen, N. F., & Li, H. (2017). Computer-assisted pronunciation training: From pronunciation scoring towards spoken language learning. In *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2016* (pp. 1–7). doi:10.1109/APSIPA.2016.7820782
- Das, R., & Sharma, U. (2016). Extracting acoustic feature vectors of South Kamrupi dialect through MFCC. In *2016 3rd International Conference on Computing for Sustainable Global Development, IN-DIACom 2016* (pp.2808–2811). Academic Press.
- Deshpande, A., Shahane, A., Gadre, D., Deshpande, M., & Joshi, P. M. (2017). A Survey of Various Chatbot Implementation Techniques. *International Journal of Computer Engineering and Applications*, 11, 7. Retrieved from [www.ijcea.com](http://www.ijcea.com)
- Fahad, S. K. A., & Yahya, A. E. (2018). Inflectional Review of Deep Learning on Natural Language Processing. In *2018 International Conference on Smart Computing and Electronic Enterprise, ICSCEE 2018*, (pp. 1–4). 10.1109/ICSCEE.2018.8538416
- Gales, M., & Young, S. (2007). The application of hidden Markov Models in speech recognition. *Foundations and Trends in Signal Processing*, 1(3), 195–304. doi:10.1561/20000000004
- Galitsky, B. (2019). Chatbot Components and Architectures. In *Developing Enterprise Chatbots* (pp. 13–47). doi:10.1007/978-3-030-04299-8\_2
- Garcia Brustenga, G., Fuertes Alpiste, M., & Molas Castells, N. (2018). *Briefing paper: chatbots in education*. doi:10.7238/elc.chatbots.2018
- González, J. (2015). Trends and Directions in Computer-Assisted Pronunciation Training. In *Investigating English Pronunciation* (pp. 314–342). doi:10.1057/978113750943
- Haristiani, N. (2019). Artificial Intelligence (AI) Chatbot as Language Learning Medium: An inquiry. *Journal of Physics: Conference Series*, 1387(1), 012020. Advance online publication. doi:10.1088/1742-6596/1387/1/012020

### **Mispronunciation Detection and Diagnosis Through a Chatbot**

Heil, C. R., Wu, J. S., Lee, J. J., & Schmidt, T. (2016). A review of mobile language learning applications: Trends, challenges and opportunities. *The EUROCALL Review*, 24(2), 32. Advance online publication. doi:10.4995/eurocall.2016.6402

Jettakul, A., Thamjarat, C., Liaowongphuthorn, K., Udomcharoenchaikit, C., Vateekul, P., & Boonkwan, P. (2018). A Comparative Study on Various Deep Learning Techniques for Thai NLP Lexical and Syntactic Tasks on Noisy Data. In *2018 15th International Joint Conference on Computer Science and Software Engineering* (pp.1–6). 10.1109/JCSSE.2018.8457368

Juang, B. H., & Rabiner, L. R. (2004). Automatic Speech Recognition – A Brief History of the Technology Development. *Elsevier Encyclopedia of Language and Linguistics*, 50(2), 637–655.

Kamath, U., Liu, J., & Whitaker, J. (2019). *Deep Learning for NLP and Speech Recognition*. doi:10.1007/978-3-030-14596-5

Kanters, S., Cucchiaroni, C., & Strik, H. (2009). The goodness of pronunciation algorithm: a detailed performance study. *Speech & Language Technology in Education -SLaTE*, (2), 2–5. Retrieved from <http://www.eee.bham.ac.uk/SLaTE2009/papers%5CSLaTE2009-33.pdf>

Khan, W., Daud, A., Nasir, J. A., & Amjad, T. (2016). A survey on the state-of-the-art machine learning models in the context of NLP. *Kuwait Journal of Science*, 43(4), 95–113.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2017). *Natural Language Processing: State of The Art, Current Trends and Challenges*. Retrieved from <https://arxiv.org/abs/1708.05148>

Leung, W. K., Liu, X., & Meng, H. (2019). CNN-RNN-CTC Based End-to-end Mispronunciation Detection and Diagnosis. In *ICASSP 2019-IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 8132–8136). 10.1109/ICASSP.2019.8682654

Li, K., Li, J., Ye, G., Zhao, R., & Gong, Y. (2019). Towards Code-switching ASR for End-to-end CTC Models. In *ICASSP 2019-IEEE International Conference on Acoustics, Speech and Signal Processing* (pp.6076–6080). 10.1109/ICASSP.2019.8683223

Li, K., Qian, X., & Meng, H. (2017). Mispronunciation Detection and Diagnosis in L2 English Speech Using Multidistribution Deep Neural Networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(1), 193–207. doi:10.1109/TASLP.2016.2621675

Li, Y., & Yang, T. (2017). Word Embedding for Understanding Natural Language : A Survey. In S. Srinivasan (Ed.), *Guide to Big Data Applications* (pp. 83–104)., doi:10.1007/978-3-319-53817-4

Liddy, E. D. (2001). Natural Language Processing. In *Encyclopedia of Library and Information Science* (2nd ed., pp. 1–15). Marcel Decker, Inc.

Luo, D., Xia, L., Zhang, C., & Wang, L. (2019). Automatic Pronunciation Evaluation in High-states English Speaking Tests Based on Deep Neural Network Models. In *2019 2nd International Conference on Artificial Intelligence and Big Data, ICAIBD* (pp.124–128). 10.1109/ICAIBD.2019.8836976

Mao, G., Su, J., Yu, S., & Luo, D. (2019). Multi-Turn Response Selection for Chatbots With Hierarchical Aggregation Network of Multi-Representation. *IEEE Access: Practical Innovations, Open Solutions*, 7, 111736–111745. doi:10.1109/ACCESS.2019.2934149

### **Mispronunciation Detection and Diagnosis Through a Chatbot**

- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. In *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings* (pp.1–12). Academic Press.
- Muangkammuen, P., Intiruk, N., & Saikaew, K. R. (2018). Automated Thai-FAQ chatbot using RNN-LSTM. In *2018 22nd International Computer Science and Engineering Conference* (pp. 1–4). 10.1109/ICSEC.2018.8712781
- Muhammad, H. Z., Nasrun, M., Setianingsih, C., & Murti, M. A. (2018). Speech recognition for English to Indonesian translator using hidden Markov model. In *2018 International Conference on Signals and Systems, ICSigSys* (pp.255–260). 10.1109/ICSIGSYS.2018.8372768
- Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, 18(5), 544–551. doi:10.1136/amiajnl-2011-000464 PMID:21846786
- Niu, C., Zhang, J., Yang, X., & Xie, Y. (2018). A study on landmark detection based on CTC and its application to pronunciation error detection. In *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)* (pp. 636–640). doi:10.1109/APSIPA.2017.8282103
- Nuruzzaman, M., & Hussain, O. K. (2018). A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks. In *2018 IEEE 15th International Conference on e-Business Engineering (ICEBE)* (pp. 54–61). 10.1109/ICEBE.2018.00019
- Panayotov, V., Chen, G., Povey, D., & Khudanpur, S. (2015). Librispeech: An ASR corpus based on public domain audio books. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5206–5210). 10.1109/ICASSP.2015.7178964
- Pellegrini, T., Fontan, L., Mauclair, J., Farinas, J., & Robert, M. (2014). The goodness of pronunciation algorithm applied to disordered speech. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp.1463–1467).
- Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., ... Vesely, K. (2011). The Kaldi Speech Recognition. *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding*.
- Qian, X., Meng, H., & Soong, F. (2012). The use of DBN-HMMs for mispronunciation detection and diagnosis in L2 english to support computer-aided pronunciation training. In *13th Annual Conference of the International Speech Communication Association 2012 INTERSPEECH*, (Vol. 1, pp.774–777).
- Roos, S. (2018). *Chatbots in education: A passing trend or a valuable pedagogical tool?* Retrieved from <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1223692&dswid=879>
- Shawar, B. A. (2017). Integrating CALL Systems with Chatbots as Conversational Partners. *Computación y Sistemas*, 21(4), 615–626. doi:10.13053/CyS-21-4-2868
- Su, P.-H., Wu, C.-H., & Lee, L.-S. (2015). A recursive dialogue game for personalized computer-aided pronunciation training. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(1), 127–141. doi:10.1109/TASLP.2014.2375572

### **Mispronunciation Detection and Diagnosis Through a Chatbot**

Thomas, N. T. (2016). An e-business chatbot using AIML and LSA. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp.2740–2742). 10.1109/ICACCI.2016.7732476

Tu, M., Grabek, A., Liss, J., & Berisha, V. (2018). Investigating the role of L1 in automatic pronunciation evaluation of L2 speech. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp.1636–1640). 10.21437/Interspeech.2018-1350

Umezawa, K. (2018). *Word2Vec: Obtain word embeddings*. Retrieved from <https://medium.com/@keisukeomezawa/word2vec-obtain-word-embeddings-885716a56270>

Wang, H., Xu, J., Ge, H., & Wang, Y. (2019). Design and implementation of an english pronunciation scoring system for pupils based on DNN-HMM. In *2019 10th International Conference on Information Technology in Medicine and Education (ITME)* (pp. 348–352). 10.1109/ITME.2019.00085

Witt, S. M., & Young, S. J. (2000). Phone-level pronunciation scoring and assessment for interactive language learning. *Speech Communication, 30*(2), 95–108. doi:10.1016/S0167-6393(99)00044-8

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine, 13*(3), 55–75. doi:10.1109/MCI.2018.2840738

Zemčík, T. (2018). A Brief History of Chatbots. *Perception, Control. Cognition*. Advance online publication. doi:10.12783/dtcse/aicae2019/31439

Zumstein, D., & Hundertmark, S. (2018). Chatbots : an interactive technology for personalized communication and transaction. *International Journal on WWW/Internet, 15*(1), 96–109.