



Facultad de Ciencias de la Administración

**Carrera de Ingeniería en Ciencias de la
Computación**

Modelo Transformers para Reducir La Ambigüedad
Sintáctica mediante Preguntas en Contextos de
Emergencia

**Trabajo de titulación previo a la obtención del
grado de Ingeniero en Ciencias de la
Computación**

Autor:

Edwin Alexander Sigüenza Sigüenza.

Director:

Marcos Patricio Orellana Cordero.

Cuenca – Ecuador

2026

DEDICATORIA

A mi mamá, Elizabeth, mi mayor apoyo y mi lugar seguro. El día a día junto a ti siempre fue tranquilo, lleno de momentos bonitos que atesoraré siempre. Gracias por tu amor incondicional, por tu paciencia en los días que no salieron tan bien, y por recordarme siempre con una sonrisa que lo bueno siempre pesa más. Eres mi mayor motivación, y este logro es tan tuyo como mío.

A mi papá, Edwin, por ser un pilar en esta etapa de mi vida. El día a día contigo fue tranquilo y disfrutable, y tu apoyo y comprensión hicieron que los momentos difíciles nunca pesaran demasiado. Gracias por tu cariño, por escucharme y por acompañarme a salir cuando necesitaba un respiro. Este logro también lleva tu nombre.

A mi hermano menor, Adrián, por ser uno de los regalos más grandes de mi vida. El abrazo que nunca falta cuando llegamos a casa, las peleas que nunca fueron en serio, los golpes de hermanos y las discusiones que siempre terminaban en risas, todo eso forma parte de mis mejores recuerdos. Enano te quiero muchísimo, y este logro lleva mucho de ti.

A mis amigos, Raúl, Andrew, Rodrigo, Roberto, Julián, John, Pablo y Juan Pablo, por hacer de estos cuatro años una experiencia que genuinamente disfruté. Hubo estrés y noches largas, pero siempre hubo más risas que preocupaciones. Gracias por los buenos momentos, los divertidos, y por convertir hasta los más difíciles en algo que hoy recordamos con una sonrisa.

AGRADECIMIENTO

A la Universidad del Azuay, por ser el espacio donde crecí no solo académicamente sino también como persona. A lo largo de estos cuatro años, cada aula, cada proyecto y cada desafío me formaron de maneras que van más allá de lo profesional.

A mi tutor de tesis, Marcos Orellana, por su guía, dedicación y paciencia a lo largo de este proceso. Su experiencia y orientación fueron fundamentales para llevar esta tesis a buen puerto.

A mi asesora metodológica, Catalina Astudillo, por su apoyo constante, su paciencia y la dedicación con la que acompañó cada etapa de este proceso. Sus orientaciones y aportaciones fueron clave para darle solidez y dirección a este trabajo.

A todos los profesores que formaron parte de mi camino durante estos cuatro años, gracias por compartir su conocimiento y experiencia con generosidad. Cada clase, cada consejo y cada enseñanza dejaron una huella que hoy se refleja en este logro.

Índice de Contenidos

DEDICATORIA.....	i
AGRADECIMIENTO	ii
Índice de Contenidos	iii
Índice de Figuras	v
Índice de Tablas.....	vi
Índice de Anexos	vii
RESUMEN	viii
ABSTRACT	viii
1. Introducción	1
2. Marco Teórico.....	4
2.1. Transformers y Procesamiento de Lenguaje Natural en Contextos de Crisis.....	4
2.2. Tipos de Ambigüedad Lingüística y su Detección	6
2.3. Estrategias de Resolución de Ambigüedad mediante Preguntas de Aclaración	7
2.4. Aplicaciones de Transformers en Dominios de Emergencia	9
2.5. Mejoras Arquitectónicas y Consideraciones de Diseño.....	10
2.6. Generación Automática de Preguntas	10
2.7. Arquitectura y Evaluación de Sistemas de Detección y Generación	11
2.8. Brecha de Investigación y Contribución Propuesta	12
3. Estado del Arte	13
3.1. Enfoques Tradicionales para Resolución de Ambigüedad Sintáctica.....	13
3.2. Transformers en Análisis de Textos de Crisis	13
3.3. Capacidades Sintácticas de los Modelos Transformer.....	14
3.4. Variantes Arquitectónicas de Transformers.....	16
3.5. Aplicaciones en Comunicación de Crisis.....	17
3.6. Sistemas de Preguntas y Respuestas para Desambiguación	17

3.7. Brechas Identificadas en la Literatura.....	19
3.8. Posicionamiento de la Investigación Propuesta.....	19
4. Métodos.....	20
4.1. Preprocesamiento de Datos.....	22
4.2. Análisis y Estandarización Lingüística.....	26
4.3. Generación Automática de Preguntas.....	28
4.4. Evaluación.....	34
5. Resultados.....	39
5.1. Detección de Ambigüedad en el Corpus.....	40
5.2. Generación de Preguntas Aclaratorias.....	43
5.3. Evaluación de Relevancia.....	45
5.4. Ejemplo de Procesamiento End-to-End.....	48
5.5. Resumen de Métricas del Sistema.....	52
6. Discusión.....	53
6.1. La Ambigüedad como Fenómeno Dominante y Sistemático.....	53
6.2. El KPI Logarítmico como Herramienta de Cuantificación.....	54
6.3. Generación de Preguntas: Volumen, Distribución y Pertinencia.....	55
6.4. Evaluación de Relevancia: Convergencia y Divergencia entre Evaluadores.....	56
6.5. Excepciones, Faltas de Correlación y Aspectos no Resueltos.....	58
6.6. Implicaciones Teóricas, Aplicaciones Prácticas y Trabajos Futuros.....	58
7. Conclusiones.....	60
8. Referencias.....	63
9. Anexos.....	67

Índice de Figuras

Figura 1 Metodología CRISP-DM.....	20
Figura 2 Diagrama Metodológico	20
Figura 3 Distribución del número de tokens por conversación y límites del filtro de selección	23
Figura 4 Distribución de conversaciones por tipo de incidente.....	24
Figura 5 Distribución de conversaciones por clave de urgencia	25
Figura 6 KPI promedio de ambigüedad por tipo de incidente.....	28
Figura 7 Distribución de conversaciones por rango de densidad de ambigüedad y tipo de incidente.....	30
Figura 8 Relación entre el KPI de densidad de ambigüedad y el score de la pregunta recomendada por tipo de incidente	41
Figura 9 Composición de señales de ambigüedad por tipo de incidente.....	42
Figura 10 Score promedio de preguntas generadas y score de pregunta recomendada por clave de urgencia	43
Figura 11 Score promedio de preguntas generadas y score de pregunta recomendada por rango de densidad de ambigüedad.....	45
Figura 12 Distribución de scores de relevancia S_i sobre el conjunto de preguntas evaluadas.	46
Figura 13 Score medio de la pregunta recomendada por tipo de incidente.....	48
Figura 14 Flujo de procesamiento end-to-end sobre la conversación 1730.....	48
Figura 15 Comparación de scores de relevancia por evaluador (score_claude y score_bert) para las siete preguntas generadas en la conversación 1730	49
Figura 16 Distribución de scores por evaluador para las siete preguntas generadas en la conversación 1730.....	51

Índice de Tablas

Tabla 1 <i>Fases de la metodología propuesta</i>	21
Tabla 2 Taxonomía de tipos de ambigüedad detectados.	27
Tabla 3 Rangos del KPI de ambigüedad y parámetros de generación.	29
Tabla 4 <i>Métricas de Evaluación</i>	38
Tabla 5 Indicadores de ambigüedad por tipo de incidente	40
Tabla 6 Distribución de señales de ambigüedad por tipo de señal y tipo de incidente ..	42
Tabla 7 <i>Indicadores por clave de urgencia</i>	43
Tabla 8 Parámetros de generación y scores de relevancia por rango KPI.....	44
Tabla 9 Métricas globales de relevancia sobre las 308 conversaciones.....	45
Tabla 10 Scores de relevancia por tipo de incidente	47
Tabla 11 Preguntas contextuales generadas y ranking de relevancia - Conversación 1730	50

Índice de Anexos

Anexo 1 <i>Manual técnico de uso</i>	67
Anexo 2 <i>Resultados conversaciones [1..50] - Ejemplo conversación 1730</i>	69

RESUMEN

La ambigüedad lingüística en las llamadas de emergencia puede retrasar la respuesta o dirigir recursos al lugar equivocado. En el ECU 911, los operadores reciben mensajes orales con estructuras incompletas, referencias vagas y sujetos omitidos. Este trabajo desarrolla un sistema que detecta señales de ambigüedad y recomienda la pregunta aclaratoria para conversación, con el objetivo de reducir la ambigüedad sintáctica en la atención de llamadas. El sistema se fundamenta en la capacidad de los modelos Transformer para representar estructuras gramaticales jerárquicas y en enfoques de generación automática de preguntas y evaluación de relevancia sin referencia anotada. La metodología siguió CRISP-DM sobre 357 conversaciones del ECU 911: análisis morfosintáctico con spaCy, generación de preguntas con GPT-OSS 20B desplegado localmente, y un esquema dual que combina el evaluador LLM-as-judge Claude Sonnet con el evaluador semántico BERT en un ranking fusionado. Sobre 308 conversaciones válidas se generaron 1.815 preguntas; la pregunta recomendada obtuvo un score medio de 0.88, frente a 0.57 del promedio del conjunto, superándolo en pertinencia operativa en mayoría de los casos. La omisión de sujeto y las referencias temporales vagas concentraron más señales detectadas en los seis tipos de incidente, confirmando que la ambigüedad es una condición estructural del dominio. El esquema dual, primera aplicación de esta combinación en emergencias en español, demostró que integrar criterio operativo con similitud semántica produce rankings más robustos que por separado, sentando las bases de un sistema que apoya al operador sin reemplazar su juicio.

Palabras clave: ambigüedad sintáctica, atención de emergencias, ECU 911, evaluación LLM-as-judge, generación de preguntas aclaratorias, modelos Transformer, procesamiento de lenguaje natural

ABSTRACT

Linguistic ambiguity in emergency calls can delay response times or misdirect resources. At ECU 911, operators receive spoken messages with incomplete structures, vague references, and omitted subjects. This study develops a system that detects ambiguity signals and recommends clarifying questions for each conversation, aiming to reduce syntactic ambiguity in call handling. The system leverages the ability of Transformer-based models to represent hierarchical grammatical structures, along with approaches for automatic question generation and reference-free relevance evaluation. The methodology follows the CRISP-DM framework on a dataset of 357 ECU 911 conversations: morphosyntactic analysis using spaCy, question generation with a locally deployed GPT-OSS 20B model, and a dual evaluation scheme combining an LLM-as-judge (Claude Sonnet) with a semantic evaluator (BERT) in a fused ranking. From 308 valid conversations, 1,815 questions were generated; the top-ranked question achieved an average score of 0.88, compared to 0.57 for the overall set, outperforming it in operational relevance in most cases. Subject omission and vague temporal references accounted for the highest number of detected signals across the six incident types, confirming ambiguity as a structural characteristic of the domain. The dual scheme—its first application in Spanish-language emergency settings—demonstrates that integrating operational judgment with semantic similarity yields more robust rankings than either approach alone, laying the groundwork for a system that supports, rather than replaces, human decision-making.

Keywords: automatic question generation, ECU 911, emergency response, LLM-as-judge evaluation, natural language processing, syntactic ambiguity, Transformer models