

Innovaciones en procesamiento de lenguaje natural (PLN) para voz

Susan Bennett, la voz original de Siri (Apple) ¹, ejemplifica cómo las voces sintéticas cada vez más realistas están transformando la interacción hombre-máquina. En los últimos años la IA conversacional integró procesamiento de voz: por ejemplo, en 2023 OpenAI anunció que ChatGPT “puede escuchar y hablar” mediante un modelo avanzado de texto-a-voz capaz de generar audio casi humano a partir de texto ². Asimismo, se apoyó en sistemas como **Whisper** (reconocimiento de voz de OpenAI) para transcribir automáticamente lo que el usuario dice ². Estas innovaciones permiten interfaces donde el usuario habla naturalmente, y el asistente entiende contexto (¡no solo comandos rígidos!). Según un reporte global, el mercado de PLN ya valía USD 24.10 mil millones en 2023 y crecerá a ~158 mil millones en 2032 (CAGR ~23%) ³. América del Norte lidera la adopción (46% del mercado en 2023 ³), y la categoría de “Respuesta de Voz Interactiva” (IVR) es una de las de mayor crecimiento ⁴. En la práctica, esto significa asistentes de voz más naturales y contextuales que hacen consultas web, gestionan tareas o incluso hablan varios idiomas (próximamente).

Esquema de reconocimiento de voz. Este diagrama simplificado ilustra cómo un sistema de **reconocimiento de voz** convierte sonidos hablados en texto. Primero el modelo acústico procesa la señal de audio en fonemas (unidades de sonido), luego un diccionario de pronunciación asocia fonemas a palabras, y finalmente un modelo de lenguaje predictivo arma oraciones coherentes (técnicas de PLN) ⁵. Innovaciones recientes –como redes neuronales profundas y grandes modelos de lenguaje– han mejorado cada etapa: la transcripción ahora alcanza alta precisión incluso con acentos o ruido, y los asistentes pueden comprender frases completas o matices emocionales. Esta base técnica ha impulsado nuevas aplicaciones de voz (p.ej., asistentes de altavoz inteligente, bots telefónicos conversacionales, generación de voz realista), ampliando el uso del PLN en sectores empresariales.

Aplicaciones por sector clave

Banca y sector financiero. En Latinoamérica la banca ha acelerado la IA conversacional para atender clientes. Por ejemplo, el banco mexicano Directo lanzó un agente de voz inteligente que atiende llamadas sin menús pregrabados: los usuarios hablan su problema libremente y el sistema contesta en lenguaje natural ⁶. Este tipo de solución reduce costos operativos y *abandono de llamadas* ⁶. En general, el ecosistema fintech latino creció ~340% entre 2017 y 2023 (de 703 a +3000 empresas) ⁷, obligando a los bancos a innovar. Estudios señalan que los chatbots tradicionales frustran al cliente (80% de usuarios informó mayor frustración ⁸) y no satisfacen sus expectativas. En contraste, los **asistentes digitales inteligentes** (ADIs) basados en PLN resuelven >80% de las consultas rutinarias ⁹ ⁸, liberando a los agentes humanos para tareas complejas. Un caso de éxito: la fintech Cyberbank Konecta (LatAm) implementó un ADI de Galileo, reduciendo en 70% las llamadas reenviadas al back-office ¹⁰ y agilizando la atención. En EE. UU. y América Latina los bancos usan IA de voz también para detección de fraude (p.ej., autenticación por biometría de voz) y asesoría financiera personalizada mediante análisis de sentimientos en llamadas grabadas.

Salud. En medicina, el PLN por voz automatiza tareas clínicas y administrativas. Un ejemplo innovador en Latinoamérica es el Hospital Internacional de Colombia: usa asistentes de voz (ej. Alexa) en quirófano para que el médico dicte fases de la cirugía en tiempo real ¹¹ ¹². Esto mejora la eficiencia del equipo (se ahorra tiempo en documentos) y permite concentrarse en el paciente. Además, se generaliza el *dictado médico* asistido por IA (reconocimiento de voz para transcribir notas clínicas), lo que ahorra horas de tipeo. El mercado regional de software de reconocimiento de voz médica creció a USD 116.2 millones en 2023 en LATAM (CAGR 11.1% pronosticado) ¹³. En EE. UU. este mercado supera los US\$1.5 mil millones ¹³. En ambos mercados, mejorar la precisión de transcripciones (incluyendo términos médicos) reduce errores de diagnóstico y costos de personal. La voz IA también se aplica en atención primaria: chatbots hablan con pacientes para pre-diagnóstico o recordatorios de medicación, complementando el triage telefónico tradicional.

Retail, comercio electrónico y bienes de consumo. El comercio aprovecha el PLN de voz para facilitar ventas y servicio al cliente. Según datos globales, en 2024 habrá ~49% de consumidores de EE. UU. usando búsqueda de voz para compras, y el 60% de compradores en línea de EE. UU. ha realizado al menos una compra mediante asistentes de voz ¹⁴. Plataformas como Amazon Alexa ya permiten ordenar productos en Walmart, Starbucks, Best Buy, etc. ¹⁵; esto impulsa la “voice commerce”. Por ejemplo, las compras por voz en todo el mundo pasaron de US\$4.6 mil millones (2021) a casi US\$20 mil millones (2023) ¹⁶ (crecimiento ~400% en dos años). En Latinoamérica algunos retailers grandes (p. ej. Cencosud) exploran asistentes conversacionales de IA para guiar la elección de productos (p. ej. comparar cafetectoras por voz) ¹⁷. Aunque el comercio por voz aún es incipiente en LatAm, las tendencias sugieren crecimiento acelerado: uno de cada cinco usuarios globales ya usa voice shopping regularmente ¹⁸. El PLN de voz también mejora atención al cliente en retail (bot conversacional por teléfono que entiende lenguaje natural, procesando reclamaciones o devoluciones sin intervención humana). La personalización basada en datos de voz ayuda a recomendar productos al cliente en tiempo real.

Turismo y hospitalidad. En este sector el PLN por voz revoluciona la atención al viajero. Se usan chatbots y asistentes multilingües de voz para reservas de vuelos, hoteles o consultas de itinerario las 24h. Un informe señala tres áreas clave: automatización de procesos repetitivos, nuevos canales de interacción (gestión eficiente de reservas, consultas o quejas) y análisis de grandes volúmenes de datos turísticos ¹⁹. Por ejemplo, cadenas hoteleras usan asistentes por voz en las habitaciones para solicitar servicios (room service por Alexa) o traducir conversaciones con huéspedes internacionales. Además, guías de voz personalizadas (apps de turismo con PLN) ofrecen información turística en el idioma del usuario. Aunque los números de LatAm son aún limitados, estos servicios atraen al turismo de negocios y ocio, mejorando la experiencia del cliente. El análisis de voz de llamadas en centros de atención al cliente también permite detectar tendencias (picos en consultas sobre vuelos/clima, etc.), ayudando a tomar decisiones operativas en aerolíneas y agencias de viajes.

Logística y otros. En logística y transporte, la IA por voz optimiza operaciones internas. Por ejemplo, empresas de transporte usan voz para guiar tareas en almacenes («pick by voice»): los operarios reciben instrucciones habladas para recoger mercancía, liberando las manos para el trabajo físico. Además, los asistentes de voz en vehículos ayudan a los conductores de reparto a navegar o reportar entregas sin apartar la vista. En México, por ejemplo, redes de transporte han experimentado con chatbots de voz para atención al cliente (seguimiento de envíos por llamada). En seguros y bienes raíces (financieras, inmobiliarias) los bots de voz responden FAQs sobre pólizas o arrendamientos, filtrando leads. Aunque no hay estudios específicos publicados para cada uno de estos subsectores, la tendencia general es clara: el

PLN de voz acelera procesos y reduce errores humanos. Según un informe de asistentes virtuales en LatAm, el mercado (chatbots y voz) fue ~USD 5.43 mil millones en 2024, creciendo ~12.9% anual ²⁰.

Costos y comparación con agentes humanos

Agentes humanos en un centro de llamadas. Tradicionalmente, los call centers dependen de equipos humanos costosos (salarios, capacitación, infraestructura). La IA de voz ofrece ahorros drásticos: por ejemplo, un caso práctico calculó que 1,538 horas/mes de agentes (a ~\$20/hora) costarían ~\$30,760 mensuales, mientras que un asistente de voz gestionando 1,000 llamadas consumiendo 2.5 horas al día (a ~\$0.25/minuto de servicio) costaría solo ~\$9,000 al mes ²¹. Es decir, se podrían ahorrar ~\$21,760 cada mes (más de \$261,000 al año) al delegar esas llamadas a la IA ²¹. De modo similar, otro análisis estima que automatizar la mitad (500/1,000) de las llamadas diarias de rutina ahorraría ~\$228,000 anuales ²². En cifras globales, Gartner prevé que la IA conversacional reducirá en US\$80 mil millones los costos de mano de obra en centros de contacto para 2026 ²³.

- **Costos de interacción:** una llamada atendida por agente humano suele costar entre ~\$5–25 (según complejidad), mientras que un agente de voz automatizado puede costar solo \$0.50–5 por llamada (solo minutos de servicio) ²¹ ²².
- **Escalabilidad:** la IA puede atender picos de llamadas simultáneamente (p. ej. cientos por segundo) sin colapsar el servicio ²⁴. Esto significa más ventas potenciales: se han reportado incrementos de hasta 3x en conversiones al reducir espera y gestionar más llamadas ²⁵.
- **Disponibilidad:** los bots 24/7 atienden fuera de horario sin turnos nocturnos ni ausencias. En sectores como salud o turismo, la atención continua puede traducirse en mayor confianza y fidelidad.
- **Calidad y cumplimiento:** las IAs pueden almacenar y aplicar inmediatamente cada actualización de políticas, evitando errores humanos y garantizando consistencia. También detectan fraudes o cumplen regulaciones de forma homogénea (p. ej. autenticación por voz).

En resumen, la **relación costo/beneficio** favorece claramente a la IA de voz para tareas repetitivas. El retorno de inversión proviene del ahorro salarial y del aumento de ingresos (más llamadas atendidas, upselling automático, retención de clientes). Además, plataformas comerciales afirman que el 49% de los consumidores usa la voz para buscar productos y el 60% ya ha comprado por ese medio ¹⁴, indicando que quedarse fuera de este canal equivale a perder ventas potenciales.

Riesgos de no implementar IA por voz

No adoptar la IA de voz implica riesgos competitivos y de negocio. Los consumidores se acostumbran a la rapidez y conveniencia de la voz: estudios muestran que el 85% de los usuarios que no reciben respuesta satisfactoria de un asistente de voz recurrirán a buscadores web ²⁶. Peor aún, el 76% de esos usuarios habría preferido una conexión fácil a un agente humano si la IA fallaba ²⁷. En la práctica, esto significa que si una empresa no dispone de un asistente de voz eficiente (o un fallback inmediato a humano), puede perder clientes frente a la competencia. Además, quedarse con tecnologías obsoletas (IVR rígido, FAQ estáticas) genera insatisfacción: por ejemplo, se reporta que 80% de clientes se frustra con chatbots anticuados ⁸.

Otros riesgos de la **no implementación** incluyen: *Incremento de costos operativos:* sin IA, la compañía debe mantener mayor plantilla para atender el mismo volumen. *Pérdida de talento y reputación:* empleados deben asumir tareas monótonas en vez de funciones estratégicas, y la marca puede ser vista como poco

innovadora. *Desventaja estratégica*: según Gartner, para 2026 las empresas líderes en IA conversacional obtendrán ahorros masivos (estimados en US\$80 mil M), por lo que no sumarse deja un enorme costo de oportunidad ²³. También, la recopilación de datos de voz propia (huella de interacciones con clientes) es valiosa para mejoras futuras y personalización; no hacerlo significa ceder esa información a terceros o perderla.

En conclusión, los datos de mercado y casos reales confirman que el PLN por voz es una tecnología madura y beneficiosa. Las empresas en banca, salud, retail, turismo o logística pueden aprovechar asistentes de voz para reducir costos, mejorar la experiencia del cliente y ganar eficiencia ⁶ ¹³ ⁹ ²⁶. Quienes ignoren esta tendencia no solo perderán ahorros significativos (hasta cientos de miles de dólares al año por centro de llamadas) ²¹ ²², sino también resiliencia competitiva en el mercado latinoamericano y estadounidense. Para una decisión informada, conviene evaluar soluciones de IA de voz probadas, alineadas con el cumplimiento local, y medir indicadores clave como reducción de tiempo de llamada, satisfacción del cliente y ROI.

Fuentes: Datos e insights de reportes e investigaciones de mercado (Fortune Business Insights ³, Grand View Research ¹³, Capital One ¹⁴), casos de uso en LATAM (Intelligent CIO ⁶, hospital en Colombia ¹¹, estudios sectoriales de América Latina ⁹ ⁷), y análisis de costo/beneficio (PolyAI ²¹, JIQ AI ²²) respaldan estas conclusiones.

¹ File:Susan Bennett VOA D0AD28DF-EAC6-49AB-8FA2-D8DC8B0CB5E0 w1023 n r0 s.jpg - Wikimedia Commons

https://commons.wikimedia.org/wiki/File:Susan_Bennett_VOA_D0AD28DF-EAC6-49AB-8FA2-D8DC8B0CB5E0_w1023_n_r0_s.jpg

² Ahora ChatGPT puede ver, escuchar y hablar | OpenAI

<https://openai.com/es-419/index/chatgpt-can-now-see-hear-and-speak/>

³ ⁴ ⁵ Tamaño del mercado, participación y crecimiento de procesamiento del lenguaje natural (PNL) [2032]

<https://www.fortunebusinessinsights.com/es/industry-reports/natural-language-processing-nlp-market-101933>

⁶ Directo launches AI voice agent to transform customer service in Mexican banking – Intelligent CIO LATAM

<https://www.intelligentcio.com/latam/2025/09/05/directo-launches-ai-voice-agent-to-transform-customer-service-in-mexican-banking/>

⁷ ⁸ ⁹ ¹⁰ IA en la Banca de América Latina: Experiencia Digital 2025

<https://www.galileo-ft.com/es/blog/ia-en-la-banca-america-latina-2025-guia/>

¹¹ ¹² Colombia transforms surgery in Latin America with voice assistants and artificial intelligence - FCV - Fundación Cardiovascular de Colombia

<https://internacional.fcv.org/en/news/33-recommended-fcv/234-colombia-transforms-surgery-in-latin-america-with-voice-assistants-and-artificial-intelligence>

¹³ Latin America Medical Speech Recognition Software Market Size & Outlook, 2030

<https://www.grandviewresearch.com/horizon/outlook/medical-speech-recognition-software-market/latin-america>

¹⁴ ¹⁵ ¹⁶ ¹⁷ The State of Voice Commerce

<https://commercetools.com/blog/the-state-of-voice-commerce>

18 Voice shopping is not a major commerce channel, but genAI and Gen Z could change that
<https://www.emarketer.com/content/voice-shopping-not-major-commerce-channel--genai-gen-z-could-change-that>

19 Position paper sobre el uso de NLP en el turismo | Observatorio IA
<https://observatorio-ametic.ai/en/publications/position-paper-sobre-el-uso-de-nlp-en-el-turismo>

20 Mercado de Asistentes Virtuales Inteligentes en América Latina, Informe 2025-2034
<https://www.informesdeexpertos.com/informes/mercado-de-asistentes-virtuales-inteligentes-en-america-latina>

21 The cost of not implementing a voice assistant - PolyAI
<https://poly.ai/blog/the-cost-of-not-implementing-a-voice-assistant/>

22 24 25 Why not adopting an AI Voice Agent could be a costly mistake
<https://jiq.ai/blog/en/why-not-adopting-an-ai-voice-agent-could-be-a-costly-mistake>

23 Compare Best AI Agents in Customer Service
<https://research.aimultiple.com/ai-agents-customer-service/>

26 27 The Rise of Voice
<https://www.invoca.com/infographics/the-rise-of-voice>