

A VOZ COMO MÍDIA DE CONHECIMENTO ORGANIZACIONAL: UMA REVISÃO INTEGRATIVA SOBRE PROSÓDIA, SEMIÓTICA E INTELIGÊNCIA ARTIFICIAL

VOICE AS A MEDIUM OF ORGANIZATIONAL KNOWLEDGE: AN INTEGRATIVE
REVIEW OF PROSODY, SEMIOTICS, AND ARTIFICIAL INTELLIGENCE

LA VOZ COMO MEDIO DE CONOCIMIENTO ORGANIZACIONAL: UNA
REVISIÓN INTEGRADORA SOBRE PROSODIA, SEMIÓTICA E INTELIGENCIA
ARTIFICIAL

Pedro Paulo da Cruz¹, Richard Perassi Luiz de Sousa², Patrícia de Sá Freire³, Solange Maria da
Silva⁴, Marily Dilamar da Silva⁵, Lucila Malagoli⁶

DOI: 10.54899/dcs.v23i87.4435

Recibido: 02/02/2026 | Aceptado: 04/02/2026 | Publicación en línea: 11/02/2026.

RESUMO

A voz humana, seja natural ou gerada artificialmente por IA, existe como mais do que uma via para a comunicação de dados; ela também estabelece um sistema complexo para a gestão do conhecimento tácito e explícito nas organizações. Devido a uma fragmentação conceitual entre os campos da comunicação, linguística e estudos de tecnologia, o artigo visa estabelecer o estado da arte atual em relação ao efeito da expressividade vocal na eficiência comunicativa organizacional. Foi realizada uma Revisão Integrativa da Literatura (RIL) utilizando o protocolo PRISMA 2020, analisando produções das bases de dados Web of Science, Scopus e EBSCO (2020–2025). Os resultados obtidos sugerem quatro eixos estruturantes: (1) a voz como indicador psicológico e preditor de liderança; (2) confiança humana em vozes geradas por IA; (3) parâmetros técnicos de fidelidade acústica; e (4) usos persuasivos para fins de marketing e gestão. Determina-se que a prosódia e a semiótica vocal são "gatilhos de conhecimento" onde a forma (entonação/ritmo) modula a interpretação do conteúdo, e as organizações precisam adquirir novas competências em "design vocal" para ambientes híbridos e mediados por algoritmos.

Palavras-chave: Comunicação Organizacional. Gestão do Conhecimento. Inteligência Artificial.

¹ Doutorando em Mídia do Conhecimento, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil. E-mail: arierom41@gmail.com Orcid: <https://orcid.org/0000-0002-0650-9233>

² Doutor em Comunicação e Semiótica, Pontifícia Universidade Católica de São Paulo, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil. E-mail: richard.perassi@ufsc.br Orcid: <https://orcid.org/0000-0003-0696-4110>

³ Doutora em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil. E-mail: patricia.sa.freire@ufsc.br Orcid: <https://orcid.org/0000-0002-9259-682X>

⁴ Doutora em Engenharia de Produção e Sistemas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil. E-mail: solange.silva@ufsc.br Orcid: <https://orcid.org/0000-0002-1925-1366>

⁵ Doutoranda em Mídia do Conhecimento, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil. E-mail: tetedila@gmail.com Orcid: <https://orcid.org/0000-0002-2322-4504>

⁶ Mestranda pelo Programa de Pós-Graduação em Engenharia, Gestão e Mídia do Conhecimento, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil.

E-mail: lucila.malagoli@prof.pmf.sc.gov.br Orcid: <https://orcid.org/0009-0006-4746-8120>

Mídia Vocal. Prosódia.

ABSTRACT

Human voice—whether natural or artificially generated by artificial intelligence—functions as more than a channel for data transmission; it also constitutes a complex system for managing tacit and explicit knowledge within organizations. Given the conceptual fragmentation across the fields of communication, linguistics, and technology studies, this article aims to establish the current state of the art regarding the effects of vocal expressiveness on organizational communication effectiveness. An Integrative Literature Review (ILR) was conducted following the PRISMA 2020 protocol, analyzing studies published between 2020 and 2025 and indexed in the Web of Science, Scopus, and EBSCO databases. The findings reveal four structuring axes: (1) voice as a psychological indicator and predictor of leadership; (2) human trust in AI-generated voices; (3) technical parameters of acoustic fidelity; and (4) persuasive uses in marketing and management contexts. The study identifies prosody and vocal semiotics as “knowledge triggers,” whereby form (intonation and rhythm) modulates content interpretation. It concludes that organizations must develop new competencies in vocal design to operate effectively in hybrid, algorithm-mediated environments.

Keywords: Organizational Communication. Knowledge Management. Artificial Intelligence. Vocal Media. Prosody.

RESUMEN

La voz humana, ya sea natural o generada artificialmente por inteligencia artificial, funciona como algo más que un canal para la transmisión de datos; también configura un sistema complejo para la gestión del conocimiento tácito y explícito en las organizaciones. Dada la fragmentación conceptual entre los campos de la comunicación, la lingüística y los estudios tecnológicos, este artículo tiene como objetivo establecer el estado actual del arte sobre los efectos de la expresividad vocal en la eficacia de la comunicación organizacional. Se llevó a cabo una Revisión Integrativa de la Literatura (RIL) siguiendo el protocolo PRISMA 2020, mediante el análisis de estudios publicados entre 2020 y 2025 e indexados en las bases de datos Web of Science, Scopus y EBSCO. Los resultados identifican cuatro ejes estructurantes: (1) la voz como indicador psicológico y predictor del liderazgo; (2) la confianza humana en voces generadas por inteligencia artificial; (3) los parámetros técnicos de la fidelidad acústica; y (4) los usos persuasivos con fines de marketing y gestión. Se concluye que la prosodia y la semiótica vocal operan como “disparadores de conocimiento”, en los que la forma (entonación y ritmo) modula la interpretación del contenido, y que las organizaciones deben desarrollar nuevas competencias en “diseño vocal” para contextos híbridos y mediados por algoritmos.

Palabras clave: Comunicación Organizacional. Gestión del Conocimiento. Inteligencia Artificial. Medios Vocales. Prosodia.



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUÇÃO

Como uma impressão digital, a voz humana é individual, e sua individualidade vem da anatomia, fisiologia e comportamento humanos. As características vocais que definem o indivíduo incluem a frequência fundamental (tom) — a altura, medida pela vibração das cordas vocais; intensidade (volume), ou força do som (de acordo com a pressão do ar e a respiração); timbre e/ou qualidade do som, derivados da ressonância da boca, nariz e garganta; duração, relacionada ao ritmo e velocidade da fala; e formantes, conjuntos de frequências que caracterizam os sons das vogais (Kreiman; Sidtis, 2011). Tudo isso constitui as características expressivas que configuram elementos ou aspectos da prosódia e influenciam a eficiência comunicacional. Dominar a qualidade do meio vocal envolve controlar tom, intensidade, timbre, velocidade, entonação e articulação do som. Portanto, a voz é também um meio ou forma de expressão, além de ser um domínio de línguas sonoras-vocalizadas, como a linguagem falada.

Na fala, a linguagem verbal é moldada em categorias de sons codificados culturalmente, que representam palavras. Assim, o desempenho comunicativo da fala baseia-se na qualidade e coerência da emissão sonora-expressiva (forma), que é como é "falada", e na seleção bem-sucedida do significado das palavras (conteúdo), que constitui "o que é falado".

Há algum tempo, houve evolução de teorias, metodologias, técnicas e tecnologia para qualificar e melhorar as propriedades de desempenho das línguas vocais, especialmente a forma da linguagem verbal falada. Recursos elétricos e eletrônicos analógicos e digitais foram projetados e criados, incluindo microfones, alto-falantes, transmissores de rádio, sistemas telefônicos, redes de comunicação online e muitos outros. Assim, além do som convencional na interação face a face, existem também sistemas tecnológicos modernos como dispositivos de amplificação ou transmissão tecnológica-digital (síncronos e assíncronos) que são projetados para aumentar a capacidade sonora ou audiovisual e a telepresença. Mas também existem sistemas artificiais de inteligência artificial (IA) digital que podem gerar ou atuar artificialmente como uma voz.

Uma vez inserida ou produzida em outros mecanismos, especialmente mídias digitais, a mídia de voz humana (gravada ou simulada) tornou-se suporte e comunicada de outras formas, ou seja, além do corpo humano. Isso é particularmente ilustrativo ao mostrar uma mídia como uma extensão do corpo, mas também como a própria mente humana, o que, claro, confirma ainda mais a identidade entre meio e mensagem e justifica a afirmação de que o conteúdo de qualquer

meio é outro meio (McLuhan, 1964).

A fala, definida como a própria linguagem, surgiu em primeiro lugar e pode ser considerada a ferramenta mais básica dentro do processo de desenvolvimento para a organização humana. Portanto, seu papel na organização social da comunidade é indispensável entre os ouvintes. Em particular, a fala é fundamental no papel de governança corporativa da comunidade, especialmente neste ambiente organizacional de ouvintes.

No domínio da gestão do conhecimento, Nonaka e Takeuchi (1995) detalham como esse processo de conversão de conhecimento tácito em conhecimento explícito (também chamado de externalização) ocorre em processos sociais e comunicativos, nos quais a linguagem e a voz desempenhadas oralmente podem atuar como pontos de gatilho para explicitar e disseminar conhecimento. Com base na perspectiva da construção de significado nas organizações, Choo (2006) afirma que o conhecimento precisa ser construído em uma série contínua de diferentes processos de interpretação ou comunicação, onde a forma como as mensagens são transmitidas (incluindo elementos vocais) molda a interpretação coletiva e as decisões. Polanyi (1966) explica que, além da fala, uma parte decisiva do conhecimento é tácita, o que é impossível de transformar completamente em linguagem formal, e, portanto, a voz é um instrumento privilegiado que medeia e transmite esse conhecimento em ambientes sociais.

A presente pesquisa é motivada pela necessidade de mapear a produção científica usando a voz em ambientes organizacionais, contornando o que atualmente está fragmentado. Em suma, é o poderoso efeito da expressão vocal que apoia a avaliação das vocalizações usadas para a medição de desempenho da comunicação falada na gestão do conhecimento organizacional.

Aqui, apresentamos o procedimento e o resultado de uma revisão integrativa da literatura (ILR) enquadrada na Declaração PRISMA 2020, com um limite determinado entre 2020 e 2025. As bases de dados digitais utilizadas foram Web of Science, Scopus e EBSCO. Além disso, alguns sites especializados tiveram pesquisas realizadas sobre temas de Comunicação.

Principais achados: Os aspectos expressivos-formais da voz humana (seja expressa, representada ou simulada) têm um impacto imediato e significativo na eficiência da fala. A prosódia é um domínio da ciência linguística. O interesse na área está focado em um conjunto de características suprasegmentais da fala, dessa forma ou melodia que significa entonação, ritmo, acento, pausas e duração que permitem a expressividade e a compreensão do que está sendo transmitido (Altman, 2012). Assim, (1) "semiótica verbal" é cognitiva, baseada no conceito de que sons são palavras. Abrange a dinâmica cognitiva que transforma estímulos acústicos em

experiências significativas (Handel, 1989) à medida que são codificados culturalmente. No entanto, os traços expressivos ou prosódicos da fala também sugerem (2) "semiótica vocal". Barthes (1990) adota essa perspectiva, onde ele observa como os modos de vocalizações da fala criam e transportam significados através da compreensão das nuances e intenções que não eram culturalmente explícitas. Além disso, Santaella (2005) fornece uma visão sobre a voz como matriz sonora, significando um dado sinal em qualidades (ícones), pistas de estados/intensões (índices) e convenções culturais (símbolos), aprimorando a leitura da vocalidade como um meio de significação. Em conclusão, com o advento de assistentes virtuais, IA e *bots* de voz, há pesquisas em andamento em relação à mediação de processos de aprendizagem através da voz e interação humano-máquina (Perri et al., 2022). Essas ideias estão fragmentadas em técnicas de prosódia e semiótica vocal através do reino da comunicação organizacional, onde são difíceis de traduzir em observações comparativas e generalizações. Mais importante, a lacuna que justifica este artigo é a fragmentação conceitual e a ausência de modelos integrativos que possam articular prosódia (forma), significação/semiótica (conteúdo e significado) e os objetivos estratégicos de comunicação organizacional, como persuasão e persuasão (Ilari, 2009).

Pergunta orientadora da revisão: Quais modelos, categorias e abordagens teóricas explicam o papel da prosódia e da semiótica vocal na construção de significado e persuasão na comunicação organizacional? Objetivo geral: descrever o estado da arte sobre prosódia, semiótica vocal e comunicação organizacional e identificar áreas de concordância teórica, lacunas e sugestões conceituais para o futuro da pesquisa.

PROCEDIMENTOS METODOLÓGICOS

Os dados foram resumidos usando o método da revisão integrativa (Botelho et al., 2011), críticas e sínteses de conhecimento foram feitas de forma integrada (Torraco, 2005), bem como o estado da arte em relação ao fenômeno estudado (Whittemore; Knafl, 2005). A técnica consiste em seis etapas: a identificação e elaboração do tema e seleção da questão de pesquisa; estabelecimento de critérios de inclusão e exclusão; identificação dos estudos selecionados; categorização dos estudos selecionados; análise e interpretação dos resultados; e apresentação da síntese do conhecimento (Botelho; Cunha; Macedo, 2011). Interpretação das etapas: (1) definição da questão norteadora; (2) elaboração da estratégia de busca (bases de dados e descritores); (3) triagem e elegibilidade (títulos, resumos e leitura completa), com eliminação de material

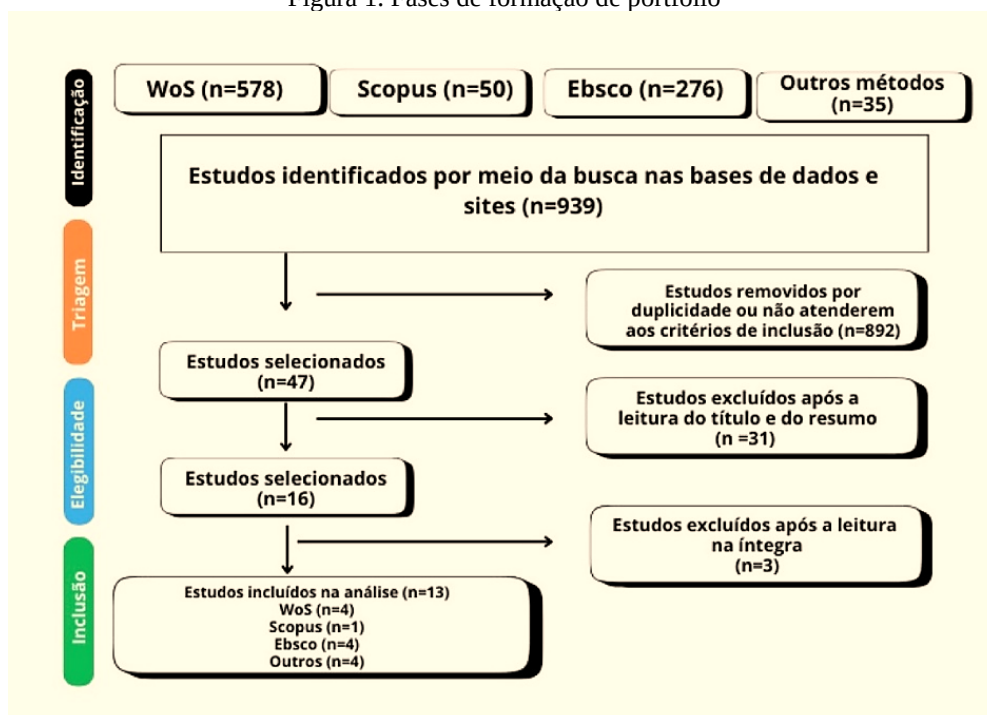
duplicado; (4) extração de dados de origem (autor, ano, contexto, foco em voz/prosódia/semiótica, método e principais achados); (5) síntese temática indutiva; e (6) formulação de proposições conceituais com base em convergências e lacunas.

Os critérios de inclusão especificados foram artigos revisados por pares publicados de 2020 a 2025 em acesso aberto (inglês, português e espanhol) e estudos empíricos e teóricos abordando o papel da voz no processo de comunicação, relacionados ou não com referência a contextos organizacionais. Documentos pagos, publicados fora do período estipulado, que indicaram conflito de interesse, em idiomas diferentes de inglês, português e espanhol, e que eram estritamente sobre saúde em oposição ao estudo da voz como mediadora do conhecimento foram excluídos desta avaliação.

Qualidade teórica da pesquisa: para informar a síntese, os estudos incluídos foram avaliados quanto a (a) clareza conceitual (prosódia/voz/semiótica), (b) correspondência entre objetivos e método, (c) transparência analítica (introduções a variáveis acústicas/prosódicas ou categorias interpretativas), e (d) relevância para contextos organizacionais (comunicação corporativa, liderança, persuasão, IA aplicada a interações).

As bases de dados foram a Web of Science, Scopus, EBSCO, combinadas com buscas complementares em sites especializados em Comunicação e bibliotecas universitárias como estratégia de busca. Os descritores corresponderam ao uso de operadores booleanos (AND/OR), com variações de idioma em inglês e português. "análise de voz" OU "comunicação vocal" OU prosódia OU "comunicação prosódica" E "comunicação organizacional" OU "pesquisa empresarial" E semiótica OU "semiótica vocal". Através de um total de 939 registros recuperados, filtros foram aplicados com base nos critérios de inclusão, removendo duplicatas, leitura de títulos e resumos, e finalmente, leitura completa, levando a 13 artigos (Figura 1).

Figura 1. Fases de formação de portfólio



Fonte: Elaborado pelos autores (2025) com base no protocolo Prisma (Page et al., 2021).

Para examinar os 13 artigos do portfólio final utilizou-se a técnica da análise temática indutiva e nível semântico, conforme descrito por Braun e Clarke (2006), que revela padrões de significado detalhando aspectos importantes dos dados em investigação. Na próxima seção apresentam-se os principais resultados.

Nota ética e transparência: este estudo é exclusivamente bibliográfico. Ferramentas de inteligência artificial generativa (por exemplo, ChatGPT) foram utilizadas apenas como apoio operacional na organização de referências, padronização textual e checagem de consistência interna, sem substituir a leitura, a interpretação e as decisões analíticas dos autores. Nenhum dado pessoal de voz foi coletado ou processado nesta etapa.

RESULTADOS

A literatura revisada exemplifica uma mudança de paradigma na análise vocal, aplicável a contextos organizacionais, de marketing e tecnológicos, na qual se move de um foco exclusivo em métricas linguísticas para a voz como uma fonte de informações emocionais, cognitivas e relacionais.

Pesquisas iniciais revelam que as qualidades da voz, como tempo, amplitude, frequência

fundamental (F0) e domínios espectrais, podem ser usadas para inferir estados internos da pessoa que fala, muitas vezes de forma mais sensível do que seu comportamento real (Hildebrand et al., 2020). Esse progresso, no entanto, traz consigo algumas implicações éticas significativas: a gravação passiva de dados vocais por um dispositivo móvel ou interface de voz, frequentemente sem a consciência dos indivíduos. Pesquisas recentes mostram que, na avaliação de personalidade e individual, padrões vocais mensuráveis externalizam algumas de suas características psicológicas. Sistemas automáticos de análise de fala surgiram como um candidato promissor para a estimativa de qualidades de personalidade, particularmente em domínios como seleção de pessoal e avaliação comportamental (Rubio et al., 2024).

Esses resultados corroboram a voz como um critério objetivo de comportamento, mas exigem robustez na captura e interpretação de dados pelos pesquisadores. Aplicações persuasivas de comunicação e marketing são as áreas de foco da análise vocal. Por exemplo, estudos anteriores mostraram que abordagens bem-sucedidas de persuasão estão correlacionadas com padrões vocais, como foco, baixo estresse e estabilidade emocional — traços que são desafiadores de avaliar por métodos tradicionais, apenas observáveis por mineração de áudio (Wang et al., 2021).

Consistente com isso, análises de vídeos de marketing sugerem que comunicações não comerciais são caracterizadas por fala mais rápida com muito pouca variação de F0, indicando que a prosódia continua a ser uma variável estratégica relativamente inexplorada na persuasão (Kuurstra, 2021). Esses achados destacam a importância de plataformas analíticas robustas para transformar som e vídeo em algo significativo para os negócios.

No campo da IA em voz, temos vários estudos indicando que o reconhecimento de emoções na fala atingiu um estágio significativo, onde os modelos estão suficientemente treinados para identificar pequenas diferenças de tom, ritmo e padrão de entonação na fala natural (Alhussein et al., 2023). No entanto, a literatura crítica destacou limitações na extração de características, na separação de emoções semanticamente próximas e em bancos de dados realistas e heterogêneos restritos, apesar do uso de aprendizado profundo (Hashem et al., 2023). Além disso, a dependência humana de sistemas de IA vocal foi encontrada como influenciada pela característica de personalidade dos especialistas, bem como pela relevância atribuída às características de entrada, ressaltando a importância de ajustar a interação humano-IA (Tutul et al., 2024).

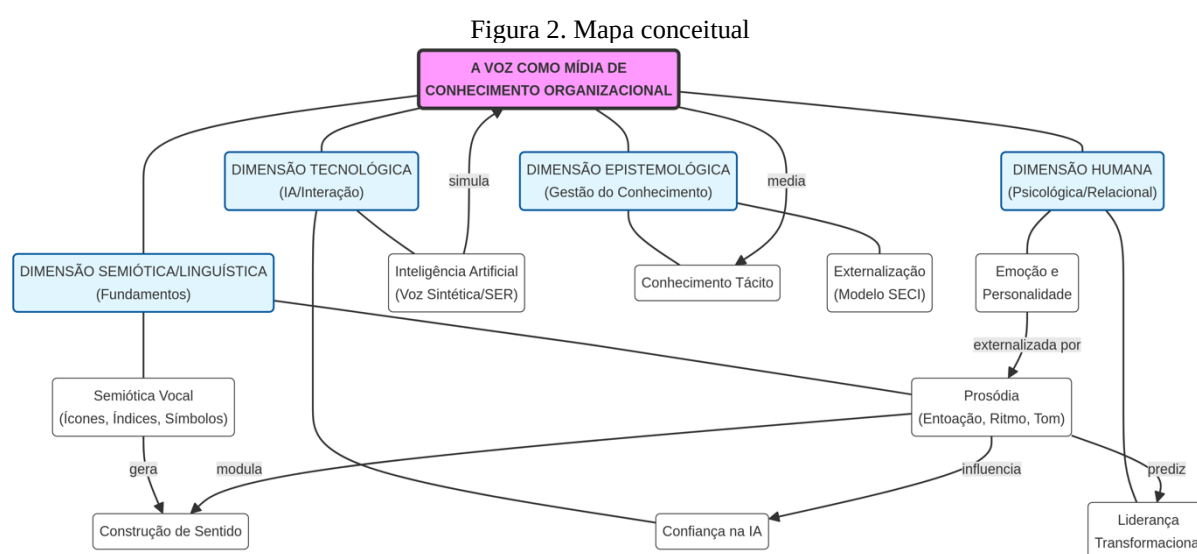
As considerações metodológicas e técnicas da captura vocal também são fatores

importantes de confiança nos resultados. Pesquisas mostram que, embora dispositivos de gravação comuns possam capturar com precisão características importantes da voz, diferenças em estratégias de amostragem podem introduzir vieses muito sérios que afetam, por exemplo, a detecção de emoções. Para garantir a qualidade analítica dos dados vocais, diretrizes explícitas são essenciais (Busquet et al., 2024).

Abordagens integrativas, no final, também fornecem uma visão mais ampla da voz como um fenômeno multimodal. Evidências recentes indicam que a frequência fundamental e a intensidade vocal são forças moduladoras de impulsos gestuais, indicando que o corpo pode ser adicionado a um sistema fonatório estendido (Pouw et al., 2020).

No contexto organizacional, análises prosódicas também mostraram relevância na percepção de liderança transformacional: por exemplo, a duração da fala e o tom vocal são indicadores-chave de carisma e influência (Hsu et al., 2025). O potencial para desenvolver a análise vocal na comunicação corporativa ainda não foi totalmente explorado, com estudos recentes apontando que a aplicação desse meio ainda está em sua infância, especialmente em áreas como desempenho, risco e fraude, exigindo mais profundidade teórica e empírica.

Como resultado, os documentos incluídos foram então organizados em quatro eixos temáticos gerais. Isso nos permitiu capturar padrões semânticos recorrentes e relações conceituais compartilhadas. Essa organização auxiliou na análise do nível semântico e significados, resultando em um diagnóstico aprimorado do campo, que é discutido na Figura 2 juntamente com os resultados resumidos na matriz de síntese apresentada a seguir.



Fonte: Elaborado pelos autores (2025).

DISCUSSÃO

Esta síntese abre caminho de uma descrição temática para uma articulação teórica, na qual a voz não é mais entendida apenas como um suporte para a linguagem verbal, mas é reconhecida como uma base simbólica, emocional e cognitiva da comunicação organizacional. A literatura revisada converge na visão de que a eficácia comunicacional não existe no conteúdo do que os participantes dizem, mas, mais criticamente, em como esse conteúdo é encenado vocalmente, nas esferas das relações sociais, relações institucionais e contextos tecnológicos.

A Dimensão Humana da Voz: Emoção, Personalidade e Liderança

Pesquisas têm mostrado consistentemente que a voz é um indicador psicológico e comportamental, refletindo externamente traços pessoais como sentimentos, afeto e traços de personalidade. Características acústicas e prosódicas permitem a estimativa de ansiedade, confiança, equilíbrio emocional, que são indicadores do estado de ser no meio da vida, tornando a vocalidade uma parte importante da comunicação na comunicação interpessoal (Hildebrand et al., 2020; Rubio et al., 2024). Esta evidência apoia a noção de que a comunicação organizacional é predominantemente mediada por métodos não verbais, através dos quais a voz é um meio privilegiado de transmitir conhecimento tácito.

De acordo com a pesquisa sobre liderança, a prosódia — particularmente a duração do discurso, entonação e modulação da voz — desempenha um papel direto na determinação de como a liderança transformacional e o carisma são percebidos. Hsu et al. (2025) argumentam que líderes que são considerados mais eficazes são aqueles que se destacam não apenas pelo discurso racionalmente fundamentado, mas também por como são ouvidos em suas reivindicações de autoridade, empatia e segurança. Este resultado corresponde perfeitamente ao conceito de *sensemaking* baseado em organização (Choo, 2006), sugerindo que a criação de significado compartilhado também é mediada por sinais de áudio que direcionam tanto interpretações, julgamentos e congruências simbólicas no nível organizacional.

Tecnologia como uma Dimensão: Inteligência Artificial, Confiança e o Problema da Voz Sintética

Esta complexidade está aumentando ainda mais com o uso crescente de inteligência artificial na comunicação organizacional. Sistemas de Reconhecimento de Emoções na Fala (SER) têm mostrado considerável capacidade de medir mudanças sutis na entonação, ritmo e intensidade da fala, inferindo estados emocionais e características psicológicas em trocas humanas facilitadas pela tecnologia (Alhussein et al., 2023; Hashem et al., 2023). A eficácia desses sistemas, no entanto, é mais do que apenas técnica, mostram os resultados. Além disso, a confiança humana em sistemas de IA vocal está fortemente correlacionada com a percepção de similaridade vocal, que é definida como tendo uma preferência por "vozes auto-similares" (Tutul et al., 2024; Jaggy et al., 2025). Assim, isso fornece a base para teorizar o surgimento de um "vale da estranheza vocal" dentro das organizações: vozes artificiais que se desviam muito dos padrões prosódicos humanos seguros tendem a fomentar desconfiança, resistência ou rejeição. Portanto, a voz sintética não é simplesmente um resultado técnico de um sistema, mas sim um aspecto semiótico que precisa ser cuidadosamente projetado para perpetuar relações de confiança, cooperação e adesão organizacional.

Voz como um Ativo de Conhecimento: Taciturnidade, Semiótica e Gestão Organizacional

Epistemologicamente, os achados adicionam evidências de que a voz é um recurso estratégico para a gestão do conhecimento. Polanyi (1966) disse em literatura anterior que "sabemos mais do que podemos dizer" e a revisão mostra que a vocalidade é uma das principais maneiras pelas quais esse conhecimento tácito é expresso. Emoções, hesitação, convicções, intencionalidade atuam de forma expressiva; são amplamente esquecidas enquanto a comunicação é limitada ao texto. Nesse sentido, os resultados correspondem diretamente ao modelo SECI de Nonaka e Takeuchi (1995), afirmando que o conhecimento é externalizado em espaços sociais na linguagem oral e na voz vocal.

A prosódia é um gatilho para converter conhecimento, moderando como o conteúdo é interpretado e moldando a formação de significado coletivo. A semiótica vocal (Santaella, 2005) permite explicar esse processo considerando a voz como um signo que funciona conjuntamente tanto sensivelmente (ícones), com estímulos emocionais e intencionais (índices), quanto

culturalmente como expectativas culturais (símbolos). Desconsiderar a dimensão vocal da comunicação organizacional, então, é ignorar uma rica constelação de dados cognitivos e afetivos que informam liderança, persuasão, tomada de decisão e confiança. A revisão mostra que a voz não apenas circula informações, mas organiza relacionamentos, legitima discursos e sustenta a organização.

Implicações para Comunicação, IA e Práticas Empresariais

Esses componentes humanos, tecnológicos e epistemológicos de forma articulada indicam a voz como o mediador mais crucial entre usuários, instituições e sistemas inteligentes. Nos setores empresariais, padrões prosódicos afetam habilidades persuasivas no marketing (Wang et al., 2021; Kuurstra, 2021), participação em reuniões e negociações, bem como a confiabilidade da instituição (Scheerschmidt et al., 2024). Simultaneamente, o nível de captura de qualidade vocal ou qualidade de gravação de voz e integridade metodológica também se tornam requisitos indispensáveis para a qualidade analítica e ética dos dados usados para informar a aplicação dessas métricas (Busquet et al., 2024; Pouw et al., 2020). Portanto, a literatura revisada apoia a criação de um Modelo Integrativo de Voz Organizacional no qual a voz é ao mesmo tempo:

- (i) um indicador psicológico e relacional;
- (ii) uma interface crucial nas relações humano-IA; e
- (iii) um canal privilegiado através do qual o conhecimento tácito pode se espalhar.

Este modelo destaca que a gestão moderna do conhecimento e a comunicação organizacional necessitam de novas competências nas áreas de design vocal, análise prosódica e conhecimento semiótico da fala, particularmente em ambientes híbridos, digitais e mediados por algoritmos.

CONSIDERAÇÕES FINAIS

Ao integrar a literatura, demonstramos que a voz é um meio primário através do qual o conhecimento é mediado, fornecendo mais do que apenas suporte para a linguagem verbal e sendo capaz de expressar dimensões cognitivas, emocionais e relacionais da comunicação humana.

Os resultados indicaram que a prosódia — o conjunto de características suprasegmentais da fala (por exemplo, entonação, ritmo, duração e intensidade) — é um fator chave na construção

do significado, clareza comunicativa e interação interpessoal, sendo crítica na determinação do desempenho de encontros organizacionais.

Não obstante tais possibilidades, a revisão da literatura mostra que existe um vazio gritante, especificamente, na literatura, em relação à voz e ao discurso organizacional, por implicação para a gestão do conhecimento e comunicação corporativa em particular. A produção científica ainda está fragmentada entre linguística, psicologia, marketing, ciência da computação, entre outros, e essa falta de continuidade no modelo teórico integrativo tem sufocado a melhoria cumulativa do assunto.

Os resultados revelam que a interseção da semiótica vocal e da análise prosódica amplia o escopo do pensamento comunicacional para incluir a transmissão de informações e a interpretação de intenção, emoção ou posição implícita nos atos de fala. Essa abordagem aplica-se especialmente a situações coletivas ou colaborativas, liderança, operação de equipe e práticas de treinamento mediadas por voz, onde o aspecto expressivo da comunicação serve como ferramenta.

As revisões também indicam que a voz serve a várias funções: como marcador psicológico, referência comportamental e como ativo estratégico nos negócios. A inteligência artificial torna-se muito atraente para a análise automática de sinais vocais para a determinação do estado emocional e padrão de comportamento nessa situação.

A partir desta revisão, foram identificadas as limitações das técnicas de captura e recuperação vocal, e a falta de um método sólido para a voz como interface em sistemas inteligentes em particular. Os achados também sugerem que as características vocais têm um impacto direto na confiança em agentes artificiais e no engajamento em interações mediadas por áudio e vídeo, por exemplo, em reuniões e comunicações corporativas. Esta evidência apoia a voz como uma entidade transversal, pois repousa na interseção de cognição, persuasão e liderança organizacional.

Como contribuição principal, esta revisão fornece uma revisão sistemática do estado da arte da pesquisa científica no campo da voz como meio de conhecimento, elucidando a convergência teórica de perspectivas, lacunas empíricas e desafios metodológicos. Os resultados fornecem evidências para apoiar a proposta de um quadro integrativo que poderia auxiliar futuras explorações nesta área para melhor moldar futuras pesquisas e práticas organizacionais para uma abordagem ao uso estratégico da voz.

Nesta linha de raciocínio, sugere-se que mais trabalhos empíricos em contexto

organizacional real sejam realizados para refletir melhor como a voz está sendo usada em relações de trabalho 'reais' e que modelos conceituais e técnicos sejam mais refinados para encorajar o uso consciente, eticamente sólido e eficaz da voz (humana e artificial) como meio de conhecimento organizacional.

REFERÊNCIAS

ALHUSSEIN, G. et al. Speech emotion recognition in conversations using artificial intelligence: a systematic review and meta-analysis. **Artificial Intelligence Review**, v. 58, n. 7, p. 1–48, 2025. Disponível em: <https://doi.org/10.1007/s10462-025-11197-8>. Acesso em: 4 maio 2025.

ALTAMAN, Cristina. Prosódia. In: CAGLIARI, Luiz Antônio (org.). **Manual de fonética acústica**. 2. ed. São Paulo: Contexto, 2012. p. 127–150.

BARTHES, Roland. **O grão da voz**: entrevistas 1962–1980. Tradução de Leyla Perrone-Moisés. São Paulo: Martins Fontes, 1990.

BOTELHO, Louise L. R.; CUNHA, Cristiano C. A. de; MACEDO, Marcelo. O método da revisão integrativa nos estudos organizacionais. **Gestão e Sociedade**, Belo Horizonte, v. 5, n. 11, p. 121–136, 2011. Disponível em: <https://doi.org/10.21171/ges.v5i11.1220>. Acesso em: 12 maio 2025.

BRAUN, Virginia; CLARKE, Victoria. Using thematic analysis in psychology. **Qualitative Research in Psychology**, v. 3, n. 2, p. 77–101, 2006. Disponível em: <https://doi.org/10.1191/1478088706qp063oa>. Acesso em: 12 maio 2025.

BRAZIL, David. **The communicative value of intonation in English**. Cambridge: Cambridge University Press, 1997. Disponível em: <https://www.cambridge.org/core/books/communicative-value-of-intonation-in-english/6A8A1C1E3A6BFA9F41A8D41D4B3A69F3>. Acesso em: 20 jan. 2026.

BUSQUET, François; EFTHYMIOU, Fotios; HILDEBRAND, Christian. Voice analytics in the wild: validity and predictive accuracy of common audio-recording devices. **Behavior Research Methods**, v. 56, p. 2114–2134, 2024. Disponível em: <https://doi.org/10.3758/s13428-023-02139-9>. Acesso em: 12 maio 2025.

CHOO, Chun Wei. **The knowing organization**: how organizations use information to construct meaning, create knowledge, and make decisions. 2. ed. Oxford: Oxford University Press, 2006. Disponível em: <https://global.oup.com/academic/product/the-knowing-organization-9780195176780>. Acesso em: 20 jan. 2026.

HANDEL, Stephen. **Listening**: an introduction to the perception of auditory events. Cambridge: MIT Press, 1989.

HASHEM, A. et al. Speech emotion recognition approaches: a systematic review. **Speech**

Communication, v. 154, art. 102974, 2023. Disponível em: <https://doi.org/10.1016/j.specom.2023.102974>. Acesso em: 12 maio 2025.

HSU, Chia-Chun; KRAJEWSKI, Jörg; FELFE, Jörg. A voz da liderança: características prosódicas como preditores da percepção de liderança. **Journal of Leadership & Organizational Studies**, v. 28, n. 1, p. 45–64, 2025. Disponível em: <https://doi.org/10.1037/mgr0000166>. Acesso em: 12 maio 2025.

ILARI, Rodolfo. **Para entender a linguagem**: ensaios sobre o cérebro e a cultura. São Paulo: Parábola Editorial, 2009.

JAGGY, Oliver; SCHWAN, Stephan; MEYERHOFF, H. S. AI-determined similarity increases likability and trustworthiness of human voices. **PLOS ONE**, v. 20, n. 3, e0318890, 2025. Disponível em: <https://doi.org/10.1371/journal.pone.0318890>. Acesso em: 12 maio 2025.

KREIMAN, Jody; SIDTIS, Diana. **Foundations of voice studies**: an interdisciplinary approach to voice production and perception. Malden: Wiley-Blackwell, 2011.

KUURSTRA, S. S. **The power of speech in persuasion**: a phonetic analysis of prosodic patterns in marketing communications. 2021.

LU, Shasha et al. Audio and visual analysis in marketing and artificial empathy. **Foundations and Trends® in Marketing**, v. 16, n. 4, p. 422–493, 2022. Disponível em: <https://doi.org/10.1561/17000000067>. Acesso em: 12 maio 2025.

MCLUHAN, Marshall. **Compreendendo a mídia**: as extensões do homem. Rio de Janeiro: Zahar, 1964.

NONAKA, Ikujiro; TAKEUCHI, Hirotaka. **The knowledge-creating company**: how Japanese companies create the dynamics of innovation. New York: Oxford University Press, 1995. Disponível em: <https://global.oup.com/academic/product/the-knowledge-creating-company-9780195092691>. Acesso em: 20 jan. 2026.

PAGE, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, London, v. 372, n. 71, 2021.

PERRI, Giulia; DI FAZIO, Luca; DI SERAFINO, Davide. Voice assistants and education: opportunities and risks for teaching and learning. **Education and Information Technologies**, v. 27, p. 1765–1782, 2022.

POLANYI, Michael. **The tacit dimension**. Chicago: University of Chicago Press, 1966. Disponível em: <https://press.uchicago.edu/ucp/books/book/chicago/T/bo6035360.html>. Acesso em: 20 jan. 2026.

POUW, Wim et al. Gesture–speech physics in fluent speech and rhythmic upper limb movements. **Annals of the New York Academy of Sciences**, v. 1491, n. 1, p. 89–105, 2021. Disponível em: <https://doi.org/10.1111/nyas.14532>. Acesso em: 4 maio 2025.

RUBIO, Víctor J. et al. Feasibility of big data voice analysis to assess personality. **Sensors**, v. 24, art. 7151, 2024. Disponível em: <https://doi.org/10.3390/s24227151>. Acesso em: 12 maio 2025.

SANTAELLA, Lucia. **Matrizes da linguagem e do pensamento**: sonora, visual, verbal. 3. ed. São Paulo: Iluminuras, 2005. Disponível em: <https://www.iluminuras.com.br/produto/matrizes-da-linguagem-e-do-pensamento/>. Acesso em: 20 jan. 2026.

SANTAELLA, Lucia. **Semiótica aplicada**. São Paulo: Thomson Learning, 2004. Disponível em: https://www.academia.edu/37504149/Semi%C3%B3tica_aplicada. Acesso em: 20 jan. 2026.

SCHEERSCHMIDT, Tim; METZLER, Dominik R. **Applications of voice analytics in corporate communication**: current state and future research directions. 2024.

TORRACO, Richard J. Writing integrative literature reviews: guidelines and examples. **Human Resource Development Review**, v. 4, n. 3, p. 356–367, 2005. Disponível em: <https://doi.org/10.1177/1534484305278283>. Acesso em: 12 maio 2025.

TUTUL, A. A.; NIRJHAR, E. H.; CHASPARI, Theodoros. Investigating trust in human–AI collaboration for speech-based data analysis tasks. **International Journal of Human-Computer Interaction**, 2024. Disponível em: <https://doi.org/10.1080/10447318.2024.2328910>. Acesso em: 12 maio 2025.

WANG, Xin (Shane) et al. Audio mining: the role of vocal tone in persuasion. **Journal of Consumer Research**, v. 48, n. 2, p. 189–211, ago. 2021. Disponível em: <https://doi.org/10.1093/jcr/ucab012>. Acesso em: 12 maio 2025.

WHITTEMORE, Robin; KNAFL, Kathleen. The integrative review: updated methodology. **Journal of Advanced Nursing**, v. 52, n. 5, p. 546–553, 2005.