

Voice Picking Software Selection Guide

This guide provides a framework for evaluating voice picking systems for use in warehouses and distribution centers.

- Learn if voice is appropriate for your warehouse and see who is using voice today.
- Understand the key benefits and features of voice solutions.
- Evaluate voice picking hardware options.
- Learn how voice technology integrates with WMS and other systems.
- Understand system pricing and implementation costs.

Table of Contents

Voice Picking 101: What is Voice Picking?	3
Beyond Voice and Beyond Picking	3
Benefits of Voice	4
What Industries Use Voice?	5
How To Decide If Voice Is Right For You	6
Software Systems Related To Voice Picking	7
Basic Voice Features	8
Human v. Synthesized Voice Direction	8
Speech Recognition Technology	8
Multimodal Support	9
Voice Hardware	10
Key Voice Software Features	11
Workflows/Processes	11
Multi-Level User Modes	12
Verification Methods	12
Exceptions	12
User Navigation and Help	12
Management And Visibility Tools	12
Deployment Options	12
Pricing and Other Costs	13
Traditional Software Licensing	13
SaaS-based Pricing	13
Transaction-based Pricing	13
Other Costs	13
More Information and Resources	14
About Lucas Systems	15

Voice Picking 101: What is Voice Picking?

Voice picking is a proven process for increasing the efficiency and accuracy of order pickers or selectors in warehouses and distribution centers. Pickers wear headsets with an industrial microphone that connects to a mobile device running the voice application.

The voice of the application tells the picker where to go and what to do, such as the warehouse pick location to go to and the quantity of the item to pick. The picker speaks into the microphone to confirm his or her work. The voice picking application includes speech recognition software that recognizes and interprets the order picker's spoken responses.

BEYOND VOICE AND BEYOND PICKING

As the phrase “voice picking” suggests, warehouses initially used these solutions to improve picking operations with voice-only technology. However, today's voice-directed applications often incorporate complementary technologies like barcode scanning or RFID. Furthermore, “voice” applications are now used to automate any manual, hands-on process in the warehouse, including receiving, put away, replenishment, cycle count, sortation, packing, cross-dock, loading, etc.

Voice picking systems are also applicable to picking and other tasks within large retail stores. For purposes of this guide, we are focusing on uses of voice within warehouses and distribution centers.



Benefits of Voice

Compared to traditional warehouse order picking processes using paper or barcode scanning (so-called RF picking, a voice-directed picking system creates a hands-free, heads-up user workflow. Rather than stopping to read, scan, write or punch keys, users listen and speak in a fluid, efficient cadence. This results in dramatic, direct pay-offs:

IMPROVE PRODUCTIVITY

Voice-Enablement: Taking paper or handheld scanners out of a worker's hand can save seconds per pick, directly improving productivity 5-10%, depending on your process (e.g., full case picking from pallets vs. piece picking from flow racks).

Voice-Optimization: Larger productivity gains – 20-90% - can be achieved by coupling voice with other process improvement techniques. Lucas customers using voice as part of our warehouse optimization solution have achieved average picking productivity gains of 36%.

REDUCE ERRORS

Using voice, pickers confirm the items and quantities they pick as they are handling items, eliminating errors and providing greater accountability versus paper-based systems. Even DCs with 99.9% picking accuracy with RF can reduce picking errors by 50% or more with voice picking.

ENSURE SAFETY

Because operators are eyes and hands free, they are more aware of surroundings. Operations that add voice solutions often see a decrease in accidents.

DECREASE TRAINING TIME

Voice-directed processes are more intuitive for users. Most Lucas customers find that a new associate can be trained and begin going solo in less than a day, and become proficient in one to two weeks.

Annual Labor Cost Savings (DC with 30 pickers)

40% Productivity Gain
\$299,985
In Annual Savings

10% Productivity Gain
\$99,445
In Annual Savings

To project the cost savings in your DC, access the Lucas online savings calculator

[Access the Savings Calculator Here](#)

Different DCs achieve vastly different results using voice, depending on a number of factors, including their approach to improving or optimizing their processes with voice. For a more in-depth discussion, see our white paper, *Three Paths To Picking Productivity With Voice and Warehouse Optimization Solutions*.

What Industries Use Voice Today?

Tens of thousands of warehouses and distribution centers around the world today use voice for outbound product distribution, direct-to-consumer ecommerce fulfillment, and for parts picking and kitting for production operations. Voice can be used for almost any product type ranging from apparel to nuts and bolts, and including items picked and shipped by weight, or length (wire coil, for example).

INDUSTRIES USING VOICE TODAY

- Retail and eCommerce Fulfillment Centers
- Beverage Wholesalers: Beer, Wine & Spirits, Soft Drinks
- Grocery and Food Distribution: Fresh, Frozen and Packaged Food
- Foodservice and Convenience
- Healthcare Products and Pharmaceuticals
- Manufacturing
- Industrial, Parts, HVAC and Electrical Supply



How To Decide If Voice Is Right For You

Voice-directed applications are used primarily in warehouses and distribution centers that are shipping products by the case or less than case quantities. DCs that only receive and ship full pallets sometimes use voice, although the productivity benefits of voice compared to other technologies (RF, or barcode scanning) may be minimal due to the large amount of travel involved in full pallet moves within a DC.

Other factors to consider, to determine if voice is a good fit for your operation:

Facility Size: Warehouses and distribution centers ranging from 50,000 ft² to 1 million ft² or larger are good candidates for voice.

Number of SKUs: The minimum number of SKUs is around 300-500. There is no upper limit as some manufacturing warehouses using voice have more than 1 million individual parts.

Number of Selectors/Pickers and Other Voice Users: Most DCs using voice have at least 10 concurrent pickers per shift. There is no upper limit to the number of users in a single facility.



Software Systems Related To Voice Picking

Voice-directed mobile applications used by warehouse workers are typically integrated with other IT and automation systems. Voice may be integrated with ERP, inventory and WMS systems either directly or indirectly using server-based voice management software or mobile work execution systems.

WAREHOUSE OPTIMIZATION SUITE

A solution that incorporates a variety of AI, machine learning and other technologies that are used to orchestrate and optimize resources - people, robots, inventory - across IT and automation systems.

ERP, WAREHOUSE MANAGEMENT SYSTEM, OR INVENTORY MANAGEMENT SYSTEM

Warehouses use ERP, WMS and custom-built systems to manage warehouse inventory, and to plan and process shipping of orders. Most WMS include RF-based receiving, picking and other applications. Voice applications may be integrated directly with a WMS or host-based inventory management system.

WAREHOUSE CONTROL SYSTEM (WCS) OR WAREHOUSE EXECUTION SYSTEM (WES)

A system that manages material handling automation systems (conveyors, sorters), typically in coordination with another inventory management or WMS system. Voice applications may be integrated directly with a WCS or WES.

LABOR MANAGEMENT SYSTEM

A system that tracks and manages labor, including time spent by task and activity. Some LMS incorporate engineered labor standards that are used to determine predicted goal times. Voice applications can provide real-time activity data to a LMS.

Basic Voice Picking

The key components of a voice system are voice application/workflow software and speech recognition software that run on a wireless mobile device. Workers wear the device on a belt (or carry it in a pocket), and most modern systems use a Bluetooth headset with a microphone that connects wirelessly to the device. The voice user interface is a dialog comprised of system voice prompts or commands, and the user's spoken responses.

SPEECH RECOGNITION TECHNOLOGY


Speaker Dependent vs. Speaker Independent Systems: Older voice systems relied on word-based recognition technology that required users to “train” every command used in the speech dialog, commonly referred to as speaker-dependent systems. Phonetic-based recognition is often used in systems that do not require users to build an individual speech model, i.e. speaker independent technology. Phonetic recognition has become the de facto standard in industrial and consumer markets as it is generally more accurate and flexible than word-based systems. Lucas uses multiple recognition technologies in a hybrid approach that combines the strengths of both technologies.

User Voice Training: Older voice systems required every new user to train every word used in the voice dialog, a process that would take from 20-40 minutes per user. Modern systems using phonetic recognition can be used with no user training, although most warehouses find it helpful to perform an initial enrollment process (2-3 minutes per user) to help users get comfortable interacting with the voice system.

Adaptive vs. Static Models: Word-based systems typically have a static user model that the recognition engine uses to interpret a user's spoken input. Recognition accuracy will often significantly degrade as users get more comfortable with the system and begin to speak more quickly. As a result, many users will need to re-record the entire model periodically. Adaptive models learn and change as users work with the system, so recognition accuracy improves over time, regardless of how a user's speech patterns may change over the course of weeks or a given shift.

HUMAN V. SYNTHESIZED VOICE DIRECTION

Voice-directed applications use either a computer-synthesized voice, or a recorded human voice for the system prompts. Recorded human voice is generally more pleasing to users and is clearer and more comprehensible at higher speeds than a synthesized voice.

 To learn more about speech recognition technology, read the Lucas white paper, [The Evolution of Speech Recognition Technology For Warehouse Operations](#).

MULTIMODAL SUPPORT

Many of today's voice-directed applications incorporate complementary technologies like barcode scanning, visual displays and RFID or NFC capabilities.

Barcode Scanning: Older voice systems exclusively used voice direction and speech recognition, but modern voice picking systems typically include the ability to scan barcodes where appropriate in the voice workflow. This is especially helpful when it is necessary to capture variable product data (lot or serial numbers, for example), and in other tasks where scanning a barcode is faster and/or more accurate than speech input.

RFID and NFC: RFID is a form of electronic tracking technology that allows a device to detect and read an electronic tag on an item or location. NFC, or near field communication, is similar to RFID, but the technology is more widely supported in smartphones and many newer voice-capable industrial computers.

Visual Displays: Old voice systems that relied on voice-only hardware required users to sign in using voice commands and buttons to scroll through a user list. Modern voice systems using smartphones or industrial mobile computers allow users to sign in on the device screen. Screen-based log in is more secure and efficient than voice log in. Device displays can also be used to provide visual information - images or lists of pick locations or items - that is more efficiently delivered on screen. The utility of visual cues is increasing as new display technologies, such as smartwatches and smartglasses, are becoming available to use with voice.

For a full discussion of the benefits of multimodal technology, see the Lucas white paper, [Warehouse Mobility Beyond Voice and RF](#).

VOICE PICKING HARDWARE

The earliest voice picking systems delivered in the 1990s and early 2000s used voice-only hardware terminals with wired headsets. Starting in 2004, the industry migrated to standard RF devices - including handheld, wearable and vehicle mount devices. These non-proprietary hardware devices were less costly and more functional than the voice-only devices. These devices all used Windows Mobile or WindowsCE operating systems and wired headsets.

In 2014, Lucas introduced voice applications that can run on Windows or Android devices, including smartphones and industrial devices. All major hardware manufacturers (Zebra, Honeywell, Datalogic) are moving to Android in their new devices, in part because Microsoft is ending support for Windows Mobile. All of these newer devices can use wireless Bluetooth headsets.

Voice-Capable Hardware Devices

Voice-Only Terminals	\$2,500+	Limited-function Windows hardware that is nearing its end of life.
Handheld RF Devices	\$1,200+	Industrial devices (Windows or Android) that provide a screen and embedded scanner; some may include a physical keypad.
Wearable RF Devices	\$2,000+	Unlike industrial handhelds, RF wearables (Windows or Android) typically do not include a scanner, as they are intended to be used with an external ring scanner.
Vehicle Mounted Devices	\$3,500+	Larger, powered devices typically include a full keyboard and scanner.
Tablets	\$300	Android or Windows tablets can be mounted on a cart or vehicle, typically used as a lower-cost alternative to a traditional industrial vehicle mount computer.
Smartphones	\$200-\$600	Android smartphones can use the camera for limited scanning or an external Bluetooth scanner.
Smartwatches	\$400	Smartwatches that can be used independent of a smartphone are now capable of supporting speech recognition applications.
Smartglasses	\$800	Smartglasses are the next frontier for picking, combining voice direction and speech recognition with augmented reality visual prompts.

Key Voice Software Features

WORKFLOWS/PROCESSES

Most distribution centers use voice for multiple picking “styles” based on product attributes, velocity, order profiles, and other factors. As noted earlier, voice can be used for a range of warehouse tasks beyond picking.

Case Picking: Popular voice picking workflows include case pick to pallet or to conveyor. Although voice picking typically eliminates the need for labels, voice systems can be used alongside labels if that is needed for delivery or other purposes. Some systems also support two-stage picking processes for slow-moving cases (sometimes referred to as PIR picking – planned inventory reserve).

Piece Picking: Popular workflows include piece picking to totes or cartons on carts, or on conveyors in a pick module. Variations include bucket-brigade pick and pass systems (zoneless picking), pick-and-put for high volume items, and two-stage picking for slow moving items.

Pallet Picking: Full pallet picking, for outbound shipping, can be interleaved with other full pallet moves (cross docking, replenishment, putaway) to reduce so-called empty travel (when a lift truck is moving from one area to another with nothing on the forks).

Receiving: Voice can be used for receiving mixed or full pallets, including verifying quantities, identifying damage, and for sorting and staging items for put-away. Barcode scanning and visual display information are commonly incorporated in the voice-directed workflow.

Putway: Voice-directed putaway can be used for full pallets, mixed pallets (multiple SKUs putaway in separate storage locations), or using carts with mixed SKUs. Barcode scanning is commonly incorporated in the voice-directed workflow, to identify/verify items for putaway.

Replenishment/Letdowns: Replenishment can be implemented for moving full pallets or mixed pallets from reserve or bulk storage to forward picking locations. In some instances, replenishment can be included in picking workflows or interleaved with other tasks.

Put Wall/Sortation: Warehouses that do not have automated sortation systems can use voice to assemble multi-line orders that have been picked in multiple zones or batches. Barcode scanning can be incorporated to identify/verify items. Items may be sorted to locations in a put wall.

Cross Dock: Various types of cross dock processes may be supported, including full transfers from receiving to shipping, or moving and sorting to outbound staging destinations.

Inventory Control/Cycle Count: In some instances, cycle count can be included in picking or other workflows, or it can be a separate task. Sophisticated voice picking systems support opportunistic cycle counts (directing a picker to count a location after he picks from it).

Truck Loading: Warehouses using voice for loading are typically loading pallets for route-stop delivery. Alternatives include system- or user-directed loading, and may include safety and other inspection steps (HACCP, for example). Systems may also support creation of a printed or electronic load map to improve delivery.

Returns: Returns are similar in some respects to receiving, with verification of quantities, inspection and disposition, and staging for putaway.

MULTI-LEVEL USER MODES

To get new users up to speed, some voice systems will offer a beginner mode that provides more information within the voice dialog. As users become more expert in the process, they may be transitioned to advanced user modes that have more concise dialogs promoting higher productivity.

VERIFICATION METHODS

Most voice picking systems use checkstrings printed on pick locations for primary verification (i.e., users speak the checkstring to confirm they are picking from the correct location). Other verification methods include voice entry of product UPCs or item numbers and scanning of product or location barcodes. Multiple verification methods may be combined in a given workflow, and different verification methods can be used selectively for different users or products.

EXCEPTIONS

All voice workflows should include means for users to identify out of stock locations or other exceptions – product damage, wrong item in slot, etc.

USER NAVIGATION AND HELP

Navigation controls allow users to skip locations or aisles, reverse pick sequences, pick base items, or to perform indirect tasks (clean up, re-stack items, etc...) Help functions – which may be presented by voice or on a display – allow workers to do their jobs without asking for assistance from supervisors, and also to avoid errors at any point in a workflow. For example, users may request additional product information (item #, picture, description, etc...) to ensure they are picking the correct item. The system should be able to provide the additional information via voice or display.

MANAGEMENT AND VISIBILITY TOOLS

Mobile voice applications capture task data that can be used by work execution and voice management systems. Voice management systems typically include basic user management and simple reporting tools, while work execution systems add work optimization and comprehensive work management tools, which providing real-time visibility and control for managers and supervisors.

Common reporting and management features include:

- User Management
- Work Management
- Productivity Reporting
- Product and Location Management
- Exceptions and Alerts
- System Configurations

DEPLOYMENT OPTIONS – CLOUD VS. ON-PREMISE

Like virtually every other software available today, voice picking systems can be deployed “in the cloud” or using an on-site server. In either architecture, the mobile voice application resides on each user’s mobile computer or smartphone, and connects to the work execution server, WMS, or other systems, over a WiFi network in the warehouse.

The majority of DCs using voice today have deployed with an on-premise server, due to the reliability and speed of communication. To address concerns about potential latency or delays with a cloud-based server or WMS, companies opting for a cloud deployment may elect to use a private cloud as part of a dedicated high-speed Wide Area Network, offering sub-second communication and greater reliability compared to public cloud infrastructure.

Pricing and Other Costs

Voice picking systems typically include software and hardware costs, plus implementation services and support. Newer pricing models, such as per-user software as a service (SaaS) and transaction-based pricing, are becoming more popular, allowing companies to avoid up-front capital investment costs by paying for software as an ongoing operational cost.

TRADITIONAL SOFTWARE LICENSING

Traditional software licensing models are based on a one-time per-user (or per mobile device) license plus a per-site or server license. These one-time license fees may decline on a per user or per-site basis with larger deployments. In addition to the annual license fees, these plans include annual software maintenance fees to provide updates and ongoing support of the software. Maintenance fees are typically in the range of 18-20% of the initial license costs.

Despite a higher initial investment, traditional licensing generally has a lower long-term cost compared to per-user service or transaction-based pricing.

SAAS-BASED PRICING (PER USER)

Cloud or on-premise voice systems can be purchased as a service, typically with a monthly fee per user paid annually. Most per-user pricing is based on a minimum three-year or five-year commitment, and the annual fees will be paid for as long as the system is in use. Companies that use the voice system beyond the initial term will usually pay more than if they had purchased the system with a traditional one-time license.

TRANSACTION-BASED PRICING

Similar to per-user pricing, DCs can pay a per-transaction service fee to use the system. Fees can be based on the number of orders, lines, or items picked or shipped, with minimum transaction levels and minimum three to-five-year terms. Companies that use the voice system beyond the initial term will usually pay more than if they had purchased the system as a traditional one-time license.

Other Costs

Mobile Devices: Warehouses and DCs may use mobile computers they are using for RF applications for voice picking, avoiding the need to purchase new computers for voice. (See the section on Voice Picking Hardware for more information.)

Wireless Network: Voice systems operate on standard WiFi networks, so DCs using RF before implementing voice may not need to make any new investments in their wireless network.

Server Hardware and System Software: For on-premise deployments, DCs will typically need to purchase a computer server with sufficient processing power and memory for the transaction volumes expected with their system. System software includes DB, security, and other utilities.

Integration, Configuration and Testing Services: The voice picking vendor will typically configure the voice applications according to the customer's workflow and other requirements, including the interface to exchange data with other systems (WMS, LMS, etc.). Those configurations and integration will be tested extensively prior to implementation.

Implementation and Training Services: On-site implementation is typically completed over the course of one or two weeks, including software installation and testing, user acceptance testing, and user training.

More Information and Resources

To understand how you can maximize productivity with voice, read [Three Paths To Picking Productivity With Voice and Warehouse Optimization Solutions](#)

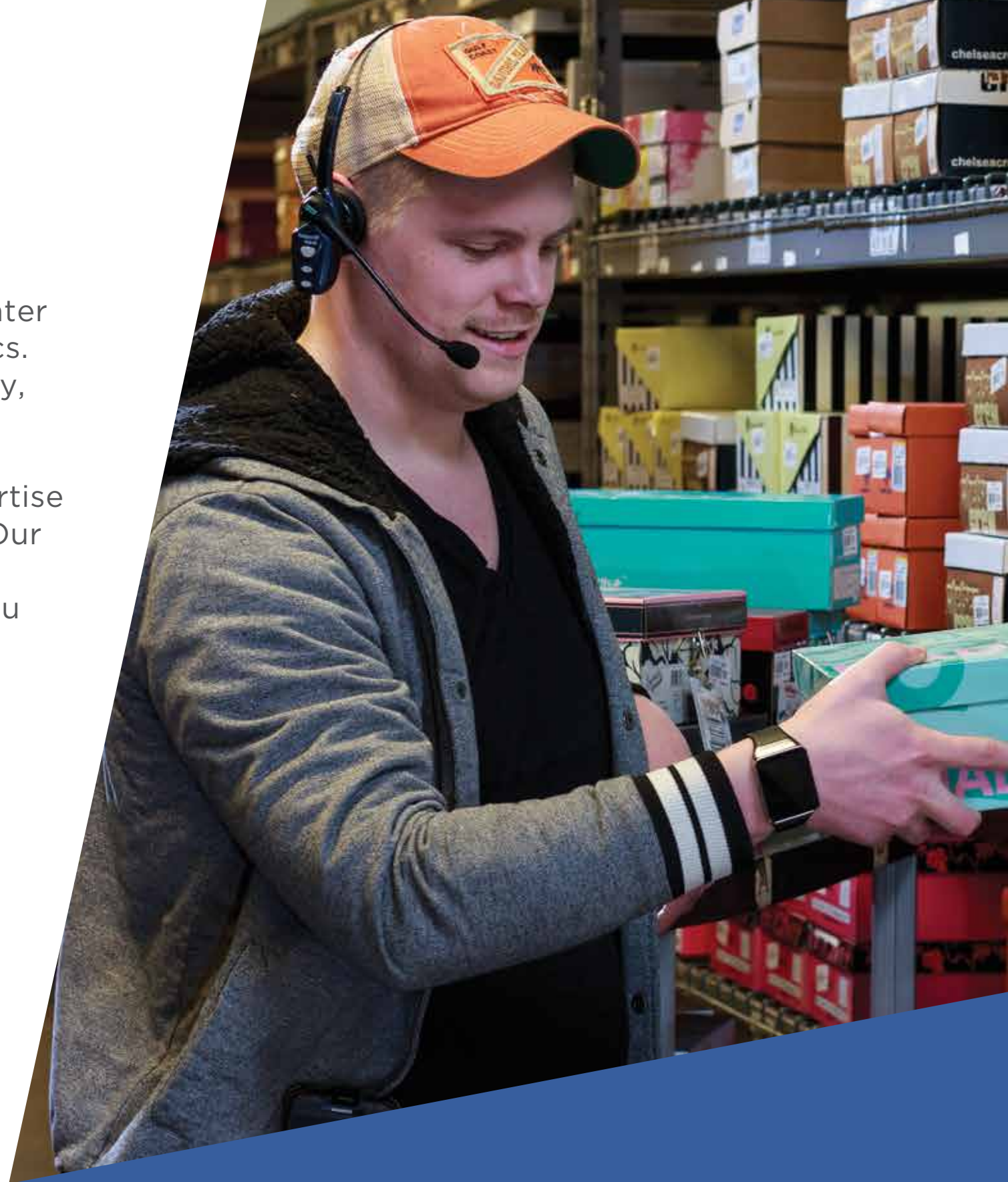
To learn more about speech recognition technology, see [The Evolution of Speech Recognition Technology For Warehouse Operations.](#)

For information on combining voice, RF and vision technology, see [Warehouse Mobility Beyond Voice and RF.](#)

About Lucas Systems

Lucas Systems helps companies transform their distribution center operations and continuously adapt to changing market dynamics. We dramatically increase worker productivity, operational agility, and accuracy and reduce the need for labor.

Lucas solutions are built on 23-plus years of deep process expertise and smart software using AI-based optimization technologies. Our solutions feature Jennifer™, the brain, voice, and orchestration engine that drives performance improvement gains. We help you make the smartest moves at the lowest cost with Jennifer™.



Lucas Systems, Inc.
150 North Meadows Drive
Wexford, PA 15090

☎ (724) 940-7000

[in](#) Lucas Systems

🌐 [Lucasware.com](https://lucasware.com)

[@Lucas_Systems](#)