THE TALKING WAREHOUSE

An eBook by Roger Byford and David Maloney

Understanding How Voice Unleashes
Higher Performance in Product Distribution

While voice-directed workflow has been a force in distribution operations for nearly two decades, it is still considered a young technology by many supply chain professionals. In a recent independent survey, only one in five supply chain managers claimed an awareness of voice. The survey also found that when distribution leaders are familiar with the benefits of voice, they overwhelmingly agree that it is an appropriate technology for distribution operations.

Supply chain and distribution managers today face a growing list of very real business challenges. They feel pressure to improve their operations, reduce costs and increase their facilities' throughputs. In a fluctuating economy, some experience the pains of growth, while others must deal with business contraction. Many have the added burden of accommodating a growing SKU base coupled with a diverse product mix that is ever-changing. And all of them are seeking to continuously improve their distribution capabilities.

Managers also must face the reality of a changing workforce. The new generation of workers will no longer have the numbers to replace aging employees as they have in the past. Older workers will need to continue on, working longer to fill the gap, and bringing with them the need to perform their jobs in less physically-demanding ways. Warehouse workplaces will continue to evolve as a melting pot with many native languages spoken. Managers will be challenged with making this new warehouse team more flexible than ever before.

To address these business challenges and to promote greater understanding of voice technology, Vocollect and DC Velocity Magazine have teamed to launch this educational series, The Talking Warehouse: Understanding How Voice Unleashes Higher Performance in Product Distribution. This partnership brings together DC Velocity, recognized by supply chain professionals as an industry voice of authority, and Vocollect, the proven authority on voice.

This educational series will help distribution and supply chain professionals gain a greater understanding of voice-directed workflow, the many ways it is applied, and the wide range of benefits companies have attained after implementing voice systems.

The series of chapter discussions will feature the basics of voice — how it was developed, how it works and where it is going. We also discuss the many ways in which voice is being deployed by leading distribution companies in a wide range of applications, including picking, receiving, replenishment, shipping, loading and inventory control.

We also will cover the infrastructure and equipment required for a successful voice implementation; making the business case for voice, including conducting the return on investment analysis; best practices using voice; the consistency voice brings to worldwide operations; and the many ways in which voice is improving operations with higher levels of productivity, increased attention on employee performance and safety, customer-centric distribution, and tools that allow managers and supervisors to better manage workflow and facility processes.

In this first chapter, we share how voice has impacted the distribution landscape. We also provide a brief history of voice and relate the story of one early adopter that recognized and capitalized on the potential of voice to revitalize its distribution operations.

We hope you'll find this educational series valuable in addressing some of the challenges you'll inevitably face in your supply chain leadership role over the coming years. Wishing you every continued success...

Sincerely,

Roger Byford
David Maloney

CHAPTER 1

The Talking Warehouse

In Chapter 1, we provided a brief historical perspective on how voice technology in distribution circles came to be, and the early experiences of Vocollect as told through the voices of its cofounders and one early adopter (Kroger).

> **Even before the dawn of recorded history, man used the spoken word to direct the work before him.**

The building of the great pyramids, the Parthenon, the Roman network of roads – all required that supervisors speak directions to their workforces to coordinate tasks. Once the electronic age was ushered in, companies relied on spoken communications, particularly the telephone, to conduct business. It is interesting to note that the first words spoken over the telephone, "Mr. Watson – come here – I want to see you," were uttered to direct a worker.

Electronic communications have since come a long way. Today computers and the Internet allow spoken words to assist in performing open heart surgery thousands of miles distant, to coordinate complex engineering feats, and to direct repairs at an orbiting space station. And our technology also makes it more efficient to get the products we use every day through the supply chain productively. Just as it was from the beginning of man's existence to this pre-Star Trek® age in which we live, there continues to be a need to communicate clearly and concisely to accomplish work.

Voice-enabled technology used in warehouses and distribution centers represents another progression in work-directed communications. Since its introduction in the distribution sector nearly 20 years ago, voice technology has dramatically enhanced the ability of thousands of leading companies across a wide range of industry verticals to achieve greater business results from their distribution processes.

All over the world, companies using voice are attaining 15-30 percent improvements in throughput, achieving 99.99 percent and higher accuracy rates, notably improving safety for their workers, and reducing new employee training times by more than 50 percent. Their customers are happier with more accurate and timely shipments. Their employees are also happier, often gaining the ability to earn greater financial rewards with stronger performance. And these organizations find that voice benefits the bottom line by providing a low total cost of ownership and a very favorable return on investment that is normally less than a year and is often as little as six months.

Yet despite its worldwide acceptance, with thousands of companies singing its praises, voice-enabled work in the distribution/supply chain arena still is not fully understood. Many supply chain professionals tell us they are familiar with voice, but few understand how it works and how it can benefit their operations.

Some may have investigated voice technology in its early years and found that its capabilities were limited. And that may have been the case two decades ago. Yet as with any emerging technology, early adopters saw enough benefits to drive voice forward to the point where all those initial concerns have since been addressed in the efficient, mature systems now available. Voice-directed work has come of age, refined through many years of research, development and practical use in distribution facilities worldwide. Today's voice systems are robust, productive and highly accurate, in every distribution environment, from grocery freezers to pharmaceuticals to apparel manufacturers. And this holds true, regardless of the spoken language and dialect.

So, what exactly is voice in the warehouse, you may ask?

In a voice-enabled warehouse, distribution workers wearing headsets and purpose-built, wearable mobile computers are directed to the tasks they must perform. It is like having someone whispering in their ears, telling them the next task to be completed.

But voice is more than just listening to directions over an earpiece – it is a system of interaction. The worker uses a microphone attached to the headset to speak back to the voice system. Recognition software within the system then analyzes their responses to determine if the work they are performing is appropriate. The result is a highly accurate and efficient system for enabling distribution work.

While voice offers many benefits to distribution center (DC) operations, speed and accuracy are the two factors that drive the attractive return on investment available with voice. Users report productivity improvements well above 20 percent, depending on the systems that voice replaces (often paper/label systems or RF scanning environments). Greater productivity means that more products can pass through a facility in a given amount of time. In some cases, the greater throughput capabilities have allowed companies to forego plans to increase the size of their buildings.

The ability of voice systems to deliver accuracy in excess of 99.99 percent also means that there are fewer returns to process, less credits to provide for those returns and happier customers who will recognize the difference voice has made in providing them with more accurate orders. Voice allows distributors to make the order correct the first time and to actually meet the demands of that elusive "Perfect Order."

GLOBAL MARKET

Today's voice technology continues to alter the distribution landscape. Vocollect, which holds the lion's share of the global voice market, accounts for more than 300,000 users worldwide. These users operate in some 60 countries and receive their voice directions in more than 35 different languages. Aggregated, these customers process more than $3 billion worth of goods to customer locations on a daily basis. So clearly, voice has impacted distribution operations in nearly every part of the world. And these are figures for Vocollect alone. Yet even with such a large user base in place, voice still is a relative unknown.

A 2008 survey by DC Velocity Magazine showed only 12 percent of respondents use voice-directed systems in any capacity. More than half – 53 percent – still are paper-based operations. More than 30 percent are picking with paper labels. These numbers align with industry estimates which show overall that voice still has only penetrated 10 percent of its potential global market. Much of this can be attributed to a lack of understanding of voice and its potential value to distribution operations. The more managers are aware of voice and its positive implications, the more likely they are to feel it is an appropriate technology for their business.

While voice technology has a multitude of applications in manufacturing, product inspection and healthcare, for our purposes, we are focusing on the use of voice in warehouse and distribution applications.

THE EMERGING ERA OF THE VOICE-ENABLED WAREHOUSE

During the past decade, many distribution processes moved from "paper-centric" to "RF-centric." Now they are moving to a "voice-centric" approach. Voice is increasingly used across many workflows, tasks and processes, and is now becoming the starting point for workflow planning. Companies are discovering that voice is work process-driven and is the technology that can best

help them optimize human/computer interactions. Full voice solutions are purpose-built to service the vast and varied potential areas of need throughout the distribution environment. Voice can be seamlessly integrated with warehouse management and ERP systems, along with additional tools required by mobile workers, including barcode scanners, printers, displays and RFID.

VOICE-DRIVEN BENEFITS

The most widely used and best established application for voice in warehouses and distribution centers is for picking, otherwise known as order selection. This is where voice first established itself in the distribution arena, and it continues to be the point of entry for most voice customers. It's usually the area where the greatest numbers of employees are positioned in the warehouse and where the company can have the most direct impact, and typically, the most immediate and measurable return on investment.

Picking is a very labor-intensive operation. It is also the place in the DC/warehouse that is highly prone to human error, as it is usually the point with the greatest amount of manual interaction. Picking relies on people to go to the right location, to select, to count and to verify – all of which are tasks subject to mistakes if not administered properly.

> **Voice meets these challenges. It keeps workers continually on task, so that there is no break in their motions and they never have to look away from their work. And since voice instructs workers step by step in their selections, it ensures top performance levels at all times.**

Voice is extremely suited for efficient picking, as workers' hands are free to perform their tasks. They do not have their hands occupied with paper, scanners, labels, pencils or anything else. Instead, their hands are readily available to pick product. Because workers are listening instead of reading a paper, label or screen, they can focus on the product being picked, which again speeds up their work for greater productivity. It also reduces the chances for errors, as they are not glancing back and forth between the paper or screen and the product location. And since their eyes are not glued to a screen or paper, workers are more aware of their surroundings, which greatly improves their safety.

In addition to picking, voice also is deployed in receiving, shipping, cycle-counting, inventory management, and many other distribution center and warehouse applications. Some facilities operate entirely on voice and have completely eliminated paper, labels, most scanning, and all other task-directed methods. Instead, voice directs and manages almost every aspect of work.

Besides directing work, voice systems also capture a wealth of data used for improving facility operations. Since voice operates in real time, there is a constant flow of accurate data generated by the system. All tasks can be tracked continuously as workers move through their assignments. The time duration of each task is measured as it is performed, allowing managers to accurately track their workforce productivity throughout the day.

In addition to determining overall productivity, voice systems give managers precise data as to which workers are more prone to errors, and who is meeting expectations and who is not. This allows managers to address problems immediately before they become customer service issues. Further, the management software's ability to provide a view on worker performance allows supervisors to identify process bottlenecks early on and other potential areas for improvement. The software helps to provide fair and accurate data for a company's incentive program, often driven by labor standards.

Associates also benefit from voice. No longer do they have to carry bulky RF scanners as they perform their work. Instead, their hands are free to focus on their tasks. Voice is easy to learn and easy to use. While it often takes weeks to train an operator to use a screen-based RF device, workers using voice can be productive within an hour and proficient within days. Training time and costs can be cut dramatically, in excess of 50 percent in many cases. Companies with a high degree of turnover or temporary labor can get new members working faster and at peak performance sooner.

Workers also prefer voice to other methods. Time and time again, workers who may have been reluctant to adopt a new technology confess they never want to go back to the old way of doing things. They cannot imagine working with anything else other than voice. They consider it a "cool" system that puts them at the cutting edge of new technology. They recognize that their companies have given them the best tools available to do their jobs, and that their employers are willing to invest in them to enhance their individual skill sets.

Voice also brings consistency to a company's operations. Not only can the same technology be used within the four walls of a facility, but it can be easily deployed at every facility within the company's network – anywhere in the world. This allows for consistent processes and easily allows benchmarking and comparison from one operation to the next. Language is also not an issue, as any language or dialect can be supported. So voice truly is a worldwide technology.

Lastly, voice is a green technology. Companies today are seeking solutions that not only save costs, but help to save the environment as well. Such green technologies allow them to be good neighbors as they perform their work. Since it eliminates huge amounts of paper and the printers and ink that go with it, voice is a socially responsible alternative.

FAVORABLE ROI

Most companies find it relatively easy to calculate return on investment with voice. It is very common for a voice system deployed in a Western nation to pay for itself within a few months. Almost every company reaches its payback within eight to 14 months. In lower-wage sectors of the world, while the ROI is not as immediate, the notable accuracy improvements, as well as the drive to attain operational excellence and consistency in their processes, provide justification for the purchase of a voice system. In this age of tightening budgets, businesses must be able to quickly see their investments produce results. Voice clearly achieves this goal. Companies also find that since voice makes their current work more productive, they can avoid the need to hire additional labor at today's ever-increasing costs of salary and benefits.

A VOICE IN THE WILDERNESS

In future chapters we will discuss in detail the many benefits of voice and how many leading companies are deploying the technology to take their distribution to the highest possible levels of productivity. But first, it would be good to learn a little of where voice came from.

The history of the voice systems used in distribution today can be traced to research and development conducted at Westinghouse Electric Corporation in Pittsburgh back in the early 1980s.

Westinghouse had established a number of research teams which worked on "intrapreneurial" pursuits. These groups were charged with developing new solutions for problems that

Westinghouse faced internally. It was also hoped that some of the solutions these teams discovered would have the potential to be marketed elsewhere.

One team was asked to explore ways to help Westinghouse's defense business record defects that occurred during the manufacture of hybrid integrated circuits. These circuits were used in the radar systems of F-16 fighter jets, which required absolute levels of quality. To inspect these circuits, workers analyzed them with microscopes – looking for damaged areas, too much solder splashed in spots, or other defects.

These inspections were hampered, though, when a worker took his eyes off the microscope to physically record the defect, location, serial number and other pertinent data. It was a longer process than it needed to be, as the inspector had to refocus each time, back and forth between paper and microscope. Work could proceed at much higher levels of productivity and accuracy if there could be a way inspectors could record their observations using voice, being prompted to provide pertinent data as they would when filling out a paper form. Such a system was soon developed by Westinghouse for these inspection efforts.

A team I (Roger) led was assigned the task of developing a production version of the voice-directed inspection system used to assist the Westinghouse defense division. I requested that Larry Sweeney, a young engineer with whom I had previously worked on a turbine project, be a member of the voice development team.

The team, which consisted of about 20 designers, eventually created an interactive voice system housed in a box about the size of a desktop personal computer. It allowed the inspectors to record their findings and to walk through a simple set of instructions and interrogative prompts.

Eventually the technology was marketed, and a few other companies purchased systems from Westinghouse to do their own inspections. While it solved the immediate needs of its initial inspection application with a person sitting at a work station, there were some problems in marketing it for other uses. The system was very expensive to produce and was not portable. It also required that inspectors be tethered to the large box which housed the system. Computing power at the time was also a limitation, as were early speech recognition systems. It would have to be reinvented in order to have any wide commercial appeal beyond its limited initial applications.

Realizing that Westinghouse was not going to devote the time and resources required for developing the technology further, Larry and I decided to strike out on our own. As engineers,

we would invent a new voice-based product designed with the portability needed for commercial appeal. Teaming with Bob Salicce, who would handle marketing and sales, we formed Vocollect in 1987.

Working in a single room above a machine shop, our young company started with a clean sheet of paper. We felt that if there was to be a market acceptance of voice, any voice product had to be user-friendly. One of Vocollect's strengths from its early years has been our ability to look at our solutions from a user's perspective and as a result, to design systems that are as ergonomically attractive as the technology allows and which provide an ease of interaction.

> "The whole idea of voice was new, risky and scary," Larry recalls. "So, we had to develop a product that was easy to operate and adapt."

We worked on creating new software and digital circuits. Among our chief developments was an application design tool which was leaps ahead of anything available at that time. We also wrote smart training software that was able to create an optimum set of usable voice patterns. Most of 1987 and 1988 were spent reinventing the product.

We found ourselves on the bleeding edge of Microsoft Windows,® starting our development with Windows 1.0 and getting rescued several times as we reached the limits of Windows by the entry of its version 2.0, then the '95 version, then NT,® and then XP.® At that time, operating systems for mobile devices were also in their infancy. We were early adopters of Windows CE,® one of just a few companies worldwide building applications with CE on devices without displays.

Adding to these constraints, early digital signal processors for processing speech were very limited – the first generation had only 128 words of program memory! It was a continual battle to make changes without adding another line to the program length. Further, our lack of CPU horsepower required adopting speech processing chips created in the UK by Marconi for military applications. As CPU power increased, we were able to move to our own algorithms implemented in software.

While progress was being made on the software, we also had to address hardware issues. We focused on three main areas to make the product mobile. First, we reduced the size and weight of the product from its desktop footprint to something that could be worn by a user. Second,

we had to create a unit that would have low power consumption, so that the battery could be relatively small and lightweight. And third, we needed to find a way of enabling the very limited microprocessors of those days to run complex speech algorithms. Our first product didn't incorporate any wireless data transmission capability. Users collected data for some period of time, and then physically connected the Talkman® voice device to a PC to transmit the data for analysis. About a year after the product's introduction, we added a so-called narrow band radio to the product to allow more frequent relay of information between the PC and the voice terminal.

Little did we know that what we were creating back then would grow from an initial foray of a voice device for a niche need into a global industry with multiple competitors headquartered around the world. With this effort, a new industry was born.

A CONNECTION IN INSPECTION

During the late 1980s and early 1990s, the commercial market for Vocollect's new voice technology products was in manufacturing inspection. Ford Motor Company was an early adopter. The automaker used the system in its manufacturing plants to record defects on the assembly line and its paint operations. Workers wearing radios and headsets would have an open dialogue with the system. They would report defects and the system would ask them a series of questions to further explain the issue.

Voice had yet to enter warehousing applications at this time, with the exception of a few minor pilot programs. The problem with the narrow band radio technology was that it could not transmit enough data to support a large number of users simultaneously, as would be required for picking and other warehouse operations. Inspection required only a small number of frequencies in comparison. But with the advent of spread spectrum technology, which companies such as Telxon were bringing from the military world into commercial products, the door opened for small test systems in warehouse operations.

Several factors made expansion into warehouses good business sense. One was that manufacturing plants, even within the same company, tended to act autonomously and have varied processes. Just because one plant could be convinced to adopt voice technology did not mean another of that manufacturer's plants would also deploy voice. The technology basically had to be

sold individually to each plant and tailored to each specific application, which often was a long, arduous process. The number of units needed for inspection was also fairly small, so there was limited potential for volume growth. And the auto industry in particular was subject to economic downturns, such as occurred in the early 1990s, as well as in 2007–2010.

Larry and I began to notice that warehouses did not offer the same marketing challenges. Grocery distribution (retail grocery, grocery distribution and third party logistics-managed distribution), our first market target for voice, represented a huge opportunity, as it is a $500 billion industry. Grocers are also more insulated from economic recessions than car manufacturers. Regardless of the economic condition of the moment, people still have to eat. They might not eat as much steak, but hamburgers and hot dogs still need to be distributed during tough times.

Second, warehousing offered the volume and scale that made voice more attractive and affordable. We could make one sale that involved hundreds of units. And distribution companies were more uniform in their operations than manufacturers. If voice was successful in one warehouse, it was likely to be adopted in a company's other distribution centers as well. And finally, refinement of spread spectrum technology allowed many users to be on the system simultaneously, as would be required for use in DC/warehouse operations.

At this point, distribution became a viable market for voice technology and the grocery business in particular became our first "beachhead" market segment. There are a number of reasons why grocery was and still is ideally suited for voice. Grocery margins are so thin that there is a huge need to maximize efficiency in order to drive profitability. Products must be turned quickly to keep logistics costs as low as possible. Most grocery products also have a limited shelf life, which further reinforces the need for speed and fast turns.

Historically, a large percentage of grocery distribution has been case-based. If workers are to handle hundreds of cases daily, it is a great advantage to have both of their hands free so they can work quickly and efficiently. It is for this reason that barcode technology was never widely adopted for grocery order-picking. It is simply too time-consuming to continually unholster, look at, scan, and holster a barcode scanner. While wrist or finger-worn scanning devices may be employed to pick cases, they are still much more cumbersome than a headset for receiving instructions and for entering data. They also are highly subject to damage, as workers reach into storage racks to pick product. As a result, grocery was, and remains, a prime market for voice.

But it certainly didn't end there. In very little time, voice found its place in markets ranging from third-party logistics and food distribution to consumer packaged goods and pharmaceuticals. For many industries, voice leveraged the same appeals of helping companies streamline their operations and improve their accuracy and throughput. For some industries, such as those dealing with products requiring lot-tracking, voice proved helpful both in improving the speed with which products got out the door and in tracing those products through the system. Retailers and others that perform a great deal of piece-picking also discovered that voice provides significant improvements in speed and efficiency in picking individual items. It also delivers the accuracy which is so crucial to their operations.

EARLY ADOPTION

When we think of pioneers, we imagine individuals of rugged character launching off into uncharted territories, facing unbelievable hardships, driving ever-onward to reach their goals.

While the pioneers of voice technology may not have wagered their lives crossing stormy seas or trudging through endless deserts, there was considerable risk in trying a new technology that had not yet proven its worth. Many placed their jobs and reputations on the line, betting voice could improve their operations, reduce their labor requirements and serve their customers better.

That is not to say it was an unmeasured risk. Each early adopter had specific needs that voice could meet. Each tested the technology under real conditions and visited other facilities that were also beginning to use the systems. Working in concert with Vocollect, these companies refined the processes, as well as the voice dialogue, to improve the functionality of the systems and to ensure that they would be successful for their intended applications.

Because picking is a large-scale process in distribution operations, the impact of using voice for picking could be seen immediately, even during the earliest deployments. An impressive ROI certainly helped with the growth of voice, especially when we could show returns in less than one year.

But still there were obstacles to overcome. One challenge we faced from the start was convincing early adopters that the system could be relied upon for 24/7 operations. Many were uncertain of a new technology from such a small, young company. In manufacturing inspection, the voice

system was not mission-critical. But in distribution, if it did not work, it could shut down the entire operation. It was up to us to prove to the market that the product worked and that we would be there to support them for years to come.

One of the earliest adopters of Vocollect Voice was The Kroger Co. (NYSE: KR), the United States' largest traditional grocery retailer, with fiscal 2009 sales of US $76.7 billion. The company was looking for a technology that addressed particular needs in its distribution operations. Kroger was, and continues to be, a visionary company looking for new ways to differentiate itself from its competitors. Kroger found success in voice, and continues to use the technology to this day to service its stores.

Kroger's venture into voice began in 1995. Doug Holmes, Kroger's Logistics Technology Manager, recalls that the Cincinnati-based grocer was in the midst of installing a new, next-generation warehouse management system (WMS).

"It was a real-time system that needed real-time data," he says. Kroger's project team looked at handheld scanners, arm-worn scanners and other systems to capture the data needed for the WMS to operate in real time, as well as to deliver picking instructions to order assemblers. "Where we got into a real problem – a show-stopper for us – was in processing our frozen foods," he recalls. "The scan guns would freeze up after about an hour of being in the cold."

Holmes says that with the new WMS, Kroger wanted to take advantage of all that real-time processing had to offer. The company did not want a system that would just dump a long list of picking instructions to a device, but one that would constantly interface with the device in real time so work could be directed continuously and updated as needed. Scanners could do this, but not if they froze. And while scanning improved accuracy, holstering and aiming proved slower than picking with the paper system previously deployed.

"Our project team toured trade shows when we were looking for our new WMS and we originally saw Vocollect there. Later, we felt this might be a fit to tackle the issues within our frozen food area," Holmes says.

Going with a very new technology was anything but a slam-dunk. "We meant to be a trailblazer with a real-time WMS, not necessarily with voice," he says. "After addressing our temperature issue, we saw voice as an opportunity to improve other operations."

The first objective was to see if Vocollect's voice unit could "stay alive" in Kroger's temperature-controlled environments. A pilot program began in the freezer area of a Kroger distribution facility in Columbus, Ohio.

There were some early concerns that Kroger wanted to address as part of the pilot. These included the dialogue between order assembler and the voice system, the response time and the battery life. Kroger worked with us to address each issue.

> **Together we built the dialogue to include the commands needed to direct work and assure accuracy, but also provide speed. Unnecessary words and other instructions were removed from the dialogue so that associates would not be slowed down. Only those prompts required to do the job were kept.**

Voice recognition systems were still very much in their infancy. Most recognition developers were creating software for use with dictation or automated phone systems.

As we will learn in the next chapter, these require a completely different set of parameters to operate, including a huge vocabulary. The early voice recognition systems were not designed for use in noisy environments such as would be found in a warehouse or distribution center. Fan noise alone in a freezer can confuse many recognition systems that are not engineered with noise cancellation capabilities. It was for this reason that we developed our own voice recognition software, designed specifically for the environment in which it would be used. Then, as now, we invested heavily in product development. Even today, over 15 percent of Vocollect's annual revenue is spent on R & D.

In addition to honing our voice recognition capabilities, we also worked to improve the response time between the Kroger WMS and the voice terminals to ensure that order assemblers were not delayed in their work while awaiting further instructions.

Batteries were another area to address. Available battery technology in 1995 was not nearly as good as can be found today. Consider the advancements made in cell phone size as an example. Cell phones in 1995 were larger for a reason. The inefficiencies of the batteries meant that larger battery cells were required to produce the needed power.

Freezers are notoriously bad environments for batteries, shortening their operating time between charges considerably. As a result, order assemblers in the freezer at Kroger had to swap out their batteries mid-shift. The frequent changing quickly wore out the battery contacts and the swing tabs that held the batteries in place on the voice terminals. Vocollect engineers beefed up the battery compartments to allow for fast and frequent changes.

"We tinkered with the hardware a lot to improve the product in those early days," recalls Holmes. "There was big involvement from Kroger's IT department to help pull this off."

The headset was another area where designers continued to make improvements. At that time, Vocollect incorporated a headset from a supplier, but this soon proved to be problematic. As users moved from a freezer environment to a humid ambient area of the DC, they would immediately find the headsets dripping in condensation that would occasionally short out the electronics. Our engineers realized that the only real way to solve this problem was to manufacture our own headsets. Today Vocollect manufactures all the headsets used in its systems, and they are all sealed to the point where the headsets can be dropped into a bucket of water without damaging the speaker or microphone.

Once Kroger installed the initial system in Columbus, Holmes says he wanted to find out first-hand what it was like to pick using voice. "After about 30 minutes of use, once I got the rhythm of the dialogue, I quickly realized that voice was actually going to be an improvement to our process and not just a solution to the temperature problem."

That original problem of finding a data collection system that would work in the freezer was solved with the voice application. But as Holmes says, voice offered more than just an ability to operate in adverse conditions.

"There are two masters to serve in distribution – speed and accuracy," he explains. The combination of the WMS and voice system enhanced accuracy over that of paper methods, as it verified the picks as they were being performed. Speed was also improved, based upon a simpler, more intuitive process. "Assemblers no longer had to read and hold paper documents; they could just listen and respond, as they move," Holmes adds. "The voice system directs activities so new associates have a quick learning curve."

In addition to speed and accuracy, the hands-free operation simply made it easier for assemblers to do their jobs, especially in the cooler where gloves were worn, making it cumbersome when trying to hold papers, peel off labels or push buttons on a scanner.

Holmes echoes the initial very favorable reaction from order assemblers in Columbus. "Even veteran associates liked the idea that some new technology was brought into their area."

Once Kroger and others found early success, other grocers soon caught on to the opportunities that voice provided. Most grocers are regionally-based and do not view their counterparts in other regions of the country as competitors. Early adopters were open to sharing their success with peers. Today, almost every major and many mid-sized grocery chains use voice for picking, and almost every one of them is a Vocollect customer.

In looking back on Kroger's experience with voice, Holmes believes that it was a good business decision. Voice solved an immediate problem Kroger was encountering; namely, picking orders using a real-time process in a freezer environment. Moving from paper and scanning to voice also improved speed and accuracy, which produced a favorable ROI. This justified the rollout of voice to other picking areas, such as refrigerated and ambient products, as well as to other facilities in Kroger's distribution network.

"It feels good to look back and realize that we were instrumental, and early, in introducing a new technology," Holmes reflects. "It proved to be a significant business opportunity for us."

SPEAKING OF RESULTS

The results in many of Vocollect's early voice installations were substantial, as voice improved productivity about 15 percent over paper and netted a whopping 35 percent improvement over scanning. Accuracy was also extremely high and much improved over paper. Larry Sweeney recalls a worker coming up to him in at one company's early pilot site saying, "I can really haul butt with this thing."

"It has been amazing to see how far we have come in reaching so many different vertical markets with voice," Larry continues. "To see an initial product with such humble beginnings spawn an entirely new industry with multiple competitors using many different types of voice devices is quite profound."

While we explored the early days of voice and its many benefits in this opening installment, we will peel back the cover a bit more in our next chapter, "Demystifying Voice." In it, we will examine the evolution and scope of speech recognition technology and how it is employed in productive ways for voice applications in the warehouse/distribution center.

CHAPTER 2

Demystifying Voice in the Warehouse

In Chapter 1, we provided a brief historical perspective on how voice technology in distribution circles came to be, and the early experiences of Vocollect as told through the voices of its cofounders and one early adopter (Kroger).

We see Chapter 2, "Demystifying Voice in the Warehouse," as a technical springboard to all the future chapters of *The Talking Warehouse*. In this chapter, we'll look at many types of voice recognition systems and explore why they are not created equally. We'll give special emphasis to voice recognition in the warehouse world, explain the optimal ways to maximize your voice recognition quality, and finally, examine the business value you can gain by using recognized best practices in speech recognition accuracy. Vocollect Adaptive Speech Recognition expertise has evolved over the course of more than two decades of research and testing and real-world application, and this chapter shares much of what has been learned through trial and error in creating and delivering voice-centric solutions to distribution organizations around the world.

While it isn't exactly "light reading," understanding the basics of speech recognition will help you fully appreciate the content of our future chapters.

> **At one time or another, we've all had to talk to a computer via a speech recognition system.**

et's be honest, this "human speaking to computer" type of experience has not always been positive. For instance, we may have phoned a customer service center and been asked by the electronic "receptionist" to state our reason for calling from a list of options. We spoke a response, and often the system failed to fully recognize what we were saying. We have to repeat the process, practically shouting at the phone and over-emphasizing our diction until the system finally understands. It leaves many of us frustrated and shaking our heads at such technology.

It is little wonder, then, that many supply chain managers fear that voice-directed workflow systems will offer similar results. If their experience with voice has been tainted by their encounters with consumer speech recognition systems, then they probably have low expectations for how well voice can succeed for their distribution operations in a "human talking to computer" type of arrangement. Little do they know that not all speech recognition systems are created equal.

In this chapter we will demystify voice and highlight its unique design to generate productivity and accuracy improvements in a distribution environment. We will pull back the curtain to reveal how there are different types of speech recognition systems that vary widely in their abilities to process information and provide desired accuracy. We will examine how some speech recognition systems have very high recognition rates in comparison to others, and how each is built to recognize various sizes of vocabularies, numbers of speakers, and types of environment (e.g., noisy versus quiet). We will look at how some speech recognition systems are built on sounds within words, while others use words themselves as the building blocks.

We will also explore why the speech recognition technology used in warehouse-based voice systems is a completely different animal from what we have experienced in the consumer arena. In particular, we will investigate the techniques the builders of warehouse voice systems use to create recognizers with very high accuracy.

The environment in which voice systems operate can greatly affect their performance. We will see how companies have engineered their products to match the demands of industrial warehouses. In this chapter we will also discuss how hardware, software and computer platforms can affect how well a speech recognition program functions.

Finally, since this book is directed to supply chain professionals, we will look at how the various pieces of the voice puzzle work together within the distribution environment to provide a system that is highly accurate in its ability to recognize a user's speech, using the attributes of speech interaction to build a productive workflow – thus driving increased business results.

This chapter focuses on how a computer understands speech from a human. To carry on a conversation, the computer obviously has to be able to speak to the human as well. We'll review computer speech output in Chapter 6 of *The Talking Warehouse*.

THE CHALLENGE

Voice recognition systems within computers have a challenge. When human beings work with a computer via screen and keyboard, we know that we are dealing with a dumb machine, so we are at least somewhat tolerant of flaws in the interaction. We also have rapid visual feedback, which in many cases helps out, e.g., "Oh. I misspelled that word – but the spelling engine in my word processor fixed it for me."

On the other hand, people are much less tolerant when it comes to spoken communications. Everything we've learned about live conversations has involved a second human being, so we fully expect the other party to respond intelligently, and often to know what we mean, even if we misspeak. And we expect this intelligence to apply in a wide variety of environments, whether it is a quiet room, a noisy coffee shop, or a loud concert.

It's little wonder, then, that users sometimes get frustrated even when a computer-based recognition system is near-perfect in understanding their speech. This becomes even more the case in a distribution environment, where every misunderstanding or slow response has an impact on productivity. Thus, it behooves the designers of warehouse voice systems to not only strive to maximize speech recognition accuracy, but also to be sensitive about managing the dialogue with the user as naturally and "intelligently" as possible.
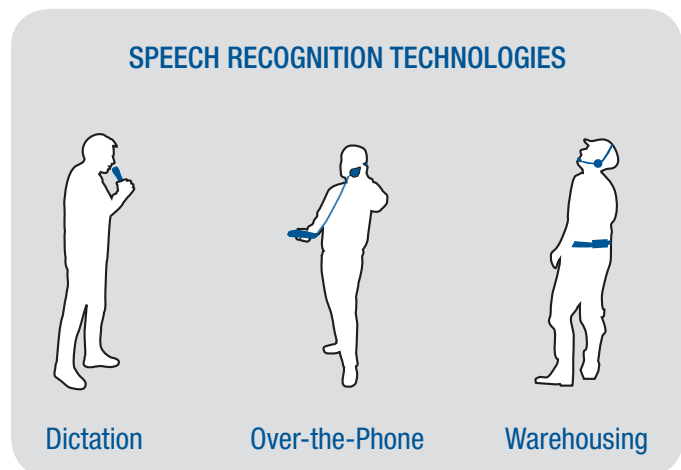
I'LL KNOW IT WHEN I HEAR IT

The goal of any type of speech recognition system is to interpret speech accurately and in a timely manner. Ideally, the system will understand everything the user wants it to hear, but will ignore everything she does not want it to hear, such as extraneous environmental noise that could be misinterpreted. Recognition systems also should be able to work in any language and within a variety of sound environments, whether loud, soft or shifting rapidly between the two extremes.

Meeting these goals can be a problem for humans, even though we typically make use of non-speech hints, such as visual information and knowledge of the subject being discussed, to help us out. For computers, which don't have this benefit, the problem is *very* challenging. How well a speech recognizer meets its objectives depends on the choice of the speech technology, how it is applied and how stringent are the user's requirements.

Let's review briefly the three most common applications of speech recognition today. Then we'll delve further into the challenges of the technology and how those challenges can be met.

DICTATION

Dictation recognizers are designed to, "Take a letter, Computer," and to immediately convert the speech to text. Some are tailored to specific domains, such as doctors or lawyers dictating specialist material. Others are simply intended to replace the keyboard for anyone writing a document on a computer.

SPEECH RECOGNITION TECHNOLOGIES

| Dictation | Over-the-Phone | Warehousing |

By its very nature, dictation software requires a huge vocabulary – literally tens of thousands of words per language. With so many words in a dictation software program's vocabulary, it is easy to see how complexity can lead to confusion. Think of English alone. How many words can be created in English simply by changing a vowel sound or the way a particular syllable is accented?

How many words are homonyms – pronounced the same, but spelled differently, with very different meanings, such as *to*, *two* and *too*? And once these various words are put into sentence structures, even more complexity is created. Furthermore, unlike systems that ask users for specific items of information, a dictation system has (almost) no idea what the user might discuss.

Fortunately, several attributes of the application make life a little easier for dictation systems. They typically function in quiet environments. Users can be expected to work with fairly high-quality microphones and to be cooperative with the recognizer, and it's acceptable to ask users to invest a little time "up-front" to familiarize the recognizer with their speech patterns (more about this process in a moment). Finally, dictation systems are usually powered from wall outlets, so they can use algorithms requiring substantial processing power and memory.

OVER-THE-PHONE

Over-the-phone recognizers are the ones most of us have encountered. They're designed to take the place, in whole or in part, of human customer service agents. Over-the-phone recognizers are typically expected to recognize a few hundred to perhaps a thousand different words – e.g., the city names in an airline information system – which is significantly fewer than in a dictation system. However, they must work for thousands, or sometimes millions, of different people with widely varying accents. Furthermore, since users do not know the limited vocabulary that these systems have or other rules designed into them, they may say unexpected things. Finally, the speech signal the system receives may be of very poor quality, coming from a noisy environment such as an airport tarmac or train station via a poorly positioned microphone and a poor phone connection. As we'll see when we compare the overall level of difficulty for the various categories of recognizer, it's not surprising that over-the-phone recognizers don't always live up to expectations.

WAREHOUSE

A warehouse recognizer doesn't have to recognize many different words – usually less than 100. However, warehouse recognizers have to operate in extremely noisy environments. They must be able to recognize words very quickly, regardless of accent or dialect, and to deal with changes in voice patterns brought about by varying noise levels (otherwise known as the Lombard effect). They must also be able to operate dependably within temperature extremes, such as working in a freezer.

SO HOW TOUGH IS IT?

One can probably begin to see already why it's reasonable to expect, and demand, much higher performance from a warehouse recognizer than from one intended to work over the phone or for dictation. Let's summarize the most important "difficulty metrics" for each of the recognizers. Then we'll discuss some of the issues in greater detail. The measures we'll use are:

Vocabulary size – How many different words must the recognizer understand?
Grammar constraints – In how many different ways can the user enter information?
Familiarity – How much does the recognizer know about the user, and vice versa?
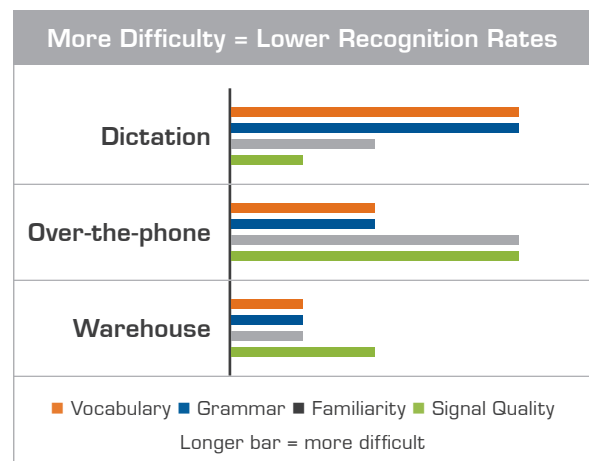**Signal quality** – What is the quality of the sound signal the recognizer receives?

Here's one view of the relative challenges:

Challenges with Speech Recognizers			
	Over-the-Phone	**Dictation**	**Warehouse**
Vocabulary size	Medium	Large	Small
Grammar constraints	Medium	Loose	Tight
Familiarity	Low	Medium	High
Signal quality	Poor	High	Medium

And here's another:

While the graph to the right is not to scale, it clearly shows that if we design a warehouse recognizer to:

- take advantage of the limited vocabulary it is required to support,

- make good use of a constrained grammar,

- benefit from knowledge of its user, and

- minimize the impact of a high-noise environment,

More Difficulty = Lower Recognition Rates

■ Vocabulary ■ Grammar ■ Familiarity ■ Signal Quality
Longer bar = more difficult

then we can make the recognizer's problem much easier to solve than is the case for either a dictation or an over-the-phone recognizer. Let's look at how designers of warehouse recognizers work with these features of the application.

PHONEMES OR WORDS?

A recognizer works by taking stored "models" of sounds and comparing them with the sounds it hears from the user. For various reasons, it's not practical for a recognizer with a big vocabulary to keep models of all the words it may hear, so such recognizers (including over-the-phone and dictation systems) use *phoneme* models instead of word models. Phonemes are the smallest sound units of a language. The word *cat*, for example, is made up of the phonemes for the hard *c*, short *a* and *t* sounds. The advantage of using phonemes for speech recognition is that there aren't many of them. The hundreds of thousands of words in the English language, for example, are all formed by combining just 44 phonemes. Unfortunately, though, the pronunciation of phonemes varies, depending on the preceding and following sounds. Add the speaker's varying inflections, accents and speech patterns, and it is easy to see why speech recognition is a lot more complex than simply hearing and interpreting.

Another challenge for phoneme-based systems is their reliance on accurate phonetic transcriptions of words, which are difficult to create. This is evident by observing text-to-speech programs, which also rely on phonetics. These systems, as their name implies, turn textual information directly into computer-generated speech.

You may have used a GPS system in your car, for instance, that uses text-to-speech to provide spoken, turn-by-turn directions. If so, then you've probably encountered times when the system attempts to speak a street name and mangles its pronunciation to the point of being rather humorous. The exceptions to pronunciation rules are many and varied, so a phoneme-based recognizer must be built with a large dictionary of phonetic spellings.

In contrast to large-vocabulary, phoneme-based systems, recognizers for small vocabularies use *word-based* models. These are designed to model the sounds of an entire word, rather than piecing together phonetic segments. There's more information in a word sound than a phoneme sound, and word pronunciation doesn't vary as much as phoneme pronunciation does, so when a small vocabulary makes them practical, word-based recognizers have a performance advantage over phoneme-based ones. Therefore, almost all distribution center and warehouse systems use word-based recognizers.

VOCABULARY SIZE AND GRAMMAR

The vocabulary, as we've discussed, is a comprehensive list of words the system can understand. It contains all the words a speaker might use in a normal conversation with the system. The vocabulary can be quite vast, so to narrow down that list and thus enable the computer to understand and respond more quickly, the system focuses on interpreting an *active vocabulary* - the possible words a speaker might use at any given moment. For instance, a telephone voice system may ask a question to which it expects only a *yes* or *no* answer. The system then has to look only at those two possible responses, which greatly increases its ability to understand what the speaker is saying, and to respond in a short time. At this point in the application, we say that the grammar is tightly constrained. Unfortunately, not all answers can be limited to "yes," "no," or other short answers from its active vocabulary. That requires the recognition system to dip deeper into its full vocabulary list – thus, the grammar becomes more loose, the chance of errors rises dramatically, and the time it takes for the computer to respond may increase.

Dictation recognizers can use a "statistical language model" to improve their accuracy. This technique makes use of grammar and context rules to help the recognizer choose between more, and less likely, sequences of words. This restricts the field of possible words, but rarely to a small and easily distinguished set. A warehouse recognizer, on the other hand, can use *tightly constrained* grammar all or most of the time. Quantity information contains only digits. Yes-no responses are fairly common. Score one point for the merits of the warehouse voice application!

FAMILIARITY – TO TRAIN OR NOT TO TRAIN?

A major difficulty with over-the-phone recognizers is that they must operate in an *untrained* mode, meaning that they don't have any information about the particular person using the system. They must be able to understand first-time callers without any history of that person to guide them. Such systems must function without knowing whether the speaker is male or female, has a high-pitched voice, or speaks in a gravelly manner.

Complicating things even more in untrained systems is that we do not all speak alike. English, for instance, sounds very different when spoken in Irish, Scottish, Canadian, South African, American and Australian accents. Even within a country, accents and dialects vary greatly. Someone from Liverpool will sound quite different than a Londoner.

To bring this idea home, think about how difficult it can be for us humans to understand someone with an unfamiliar accent. At first, we may pick up only a few words. After listening for a couple of minutes, our comprehension rate rises. If the speaker continues to discuss the same subject, reusing words with unusual pronunciations, we eventually figure out those words, and our understanding increases significantly. What if a recognizer were given the opportunity to learn from its users by hearing them say, in advance, the words they will use as they work with the system? We'd expect the computer to benefit in recognition rate and response time, just as we do. And it does.

Warehouse recognizers can be "fully trained" to better understand the user, which greatly increases recognition rates. A user simply has to invest 10 to 20 minutes teaching the system how she pronounces the words used in the warehouse environment. This is done prior to using the system on the facility floor. The ability for the system to learn from the user greatly affects overall recognition and is, as we will see, a smart investment of an employee's time.

The system prompts the user to speak the words in the vocabulary, which she repeats back into the headset's microphone. The software registers these words as typical responses from her, determining her unique tonal pitch, accent, inflections and pronunciations of each word. The recognizer uses the responses to build speaker-specific models of all the words in the vocabulary. It's easy to see that having the opportunity to record and work with each user's individual voice patterns makes the recognizer's task far easier. A trained recognizer is key to achieving the high levels of performance offered by today's industrial voice systems. Another benefit is that, because a trained recognizer knows exactly how its current user speaks, it can dramatically narrow down the range of sounds it will interpret as the word "one" (for example). In noisy environments, that makes such a recognizer much less prone to interpreting background noise or "off-topic" speech as an intended utterance.

Additionally, since a trained system bases its recognition on a person's individually-recorded voice patterns, it does not have to consider how the word *should* be pronounced, only how the worker *does* pronounce the word. It no longer relies on what the typical pronunciation should be, as dictation systems would. For instance, let's look at how the system determines the number "1" has been spoken to it. When the worker trained the recognizer, she was asked to say "1." The recognizer assumes that whatever sound she made at that time represents the number 1, and it stores that sound for future comparison with her speech. Trained voice recognizers, therefore, work equally well without modification for any language or dialect.

An untrained recognizer, on the other hand, must be specifically programmed for each language it is required to recognize, and must be told ahead of time what language the speaker will be using. In the warehouse environment where multilingual workforces are becoming commonplace, workers using a fully trained recognizer can speak to the system in whatever language makes them most comfortable. Such flexibility greatly increases the ease with which workers adapt to a voice system.

We also have to consider how familiar the user is with the system he's talking to. If not familiar with it, he's likely to say things the recognizer doesn't expect. Even if the recognizer determines that the user said something that was outside the expected vocabulary, about all it can do is to either ignore the speech or apologize for not understanding it. Neither is very helpful to the user. This is an unavoidable problem with over-the-phone systems. It's less of a challenge for dictation systems, both because they are designed to respond to a very large vocabulary and loose grammar, and (for most users) visual feedback can help to clear up any confusion.

Industrial voice users work with their systems for many hours per day, conducting a relatively limited range of transactions. The associates quickly learn, either through classroom or on-the-job training, what responses are acceptable to the recognizer at each point in their dialogue with it. Unlike most users of over-the-phone systems, industrial recognizer users generally have a strong vested interest in being successful with the system and don't have the option of pressing zero to reach a human being.

The fact that the users become familiar with the system through formal training and extended use adds to the performance difference between these specialized systems and more general-purpose, over-the-phone and dictation recognizers.

Increased performance is good, but one has to ask if the cost of training a system pays for itself by improving recognition accuracy. Vocollect has evaluated the return on investment associated with trained systems in comparison to untrained systems. We first calculated the cost to train a warehouse recognizer, a process which takes 10-20 minutes. At US $30 per hour for salary and benefits, training the recognizer costs a facility about US $10 in employee time. We compared that expense with the costs the same employee would incur if he or she experienced the higher recognition error rates associated with untrained systems. The criteria for determining the additional costs were based on delays in work time caused by the need to repeat words or add extra words to allow the untrained system to process a response correctly.

According to the study's calculations, if training the recognizer reduces the error rate by one percentage point (from 2% to 1%, for example), the payback period for the investment in pre-use training (recording the reference word sounds) is less than six work days. And the ongoing cost savings per device each year is roughly US $450 for case-picking and US $2400 for piece-picking. In a two-shift operation, the annual savings per device doubles. Our head-to-head tests of several untrained warehouse recognizers suggest that the actual performance difference is likely to be several percentage points. Add up the total number of devices used in a facility, and it is easy to see that the savings a word-based, fully trained recognizer provides are substantial.

ICH SPRECHE DEUTSCH

A fully trained recognizer brings an additional benefit for multilingual workforces. In many parts of the world, a large percentage of warehouse workers are non-native. This is certainly the case in North America, where many workers speak English as their second language. In European Union countries, workers easily cross borders to work – a typical warehouse may have five or more native languages represented. Untrained systems, therefore, must be able to recognize responses laden with strong accents. But how much does an accent affect recognition accuracy?

Vocollect has conducted considerable research into this area. In a recent study, technicians ran tests comparing the performance of our own trained speech recognizer against several untrained recognizers available for the warehouse. The modeling for our study was based on non-native workers speaking English with moderate accents. The results showed that, for distribution usage, the increase in word error rate when moving from a trained recognizer to an untrained one is likely to be several percent or more. In fact, for speakers with moderate-to-strong accents, the untrained recognizers showed error rates *ranging from 6% to more than 20%,* while the trained recognizer's error rate on the same task was below 1%. This is because a trained recognizer is not affected by accent. The system simply learns how the worker pronounces the words – accent and all. Accent is no longer a factor.

In addition, the ability for workers to speak to the recognizer in the language of their choice, regardless of accent, improves job satisfaction and virtually guarantees that all workers in a facility will be able to use the voice system to its best effect. Enhancing the performance and satisfaction of a multilingual environment is, therefore, another benefit of deploying a trained recognizer in a distribution center or warehouse.

MANAGING NOISE: THE HEADSET MATTERS

We stated earlier that a challenge for warehouse recognizers is minimizing the impact on performance of the noisy environment in a distribution center. The background noise from lift trucks, fans, blaring music, public address systems and other equipment makes it challenging to have exacting speech recognition every time. Even more challenging for the voice equipment designer, warehouse systems must work in freezers, on sun-baked loading docks, and everywhere in between. While the design and implementation of the recognizer itself plays a big part in maximizing performance amid noise, it's also crucial to provide the recognizer with the best possible input, maximizing the so-called signal- (user speech) to-noise (other sounds) ratio. Vocollect designs or directs the design of all of our own headsets, and we do this for a reason. Off-the-shelf headsets are simply not optimized to work for rugged material-handling applications.

A number of factors determine headset performance and ultimately affect the accuracy of speech recognition. For example, the microphone must have good noise cancellation characteristics, eliminating noises that originate far from the microphone, while capturing sound that comes from the user speaking directly into the microphone. Noise-canceling microphones require very tight manufacturing tolerances, making them expensive; however, a poor microphone will compromise the performance of the entire system. Unfortunately, the response of a noise-canceling microphone varies widely if it is moved, relative to the user's mouth, by even a small distance. This means that the recognizer must be designed to compensate, as best it can, for any such variations. The headset should be designed to make it easy and intuitive to position the microphone the same way, all day and every day.

The microphone must also perform in extreme conditions, even after being subjected to dirt, debris, and liquids common to distribution environments. The microphone and terminal circuitry have to work together and must be built with an adequate *dynamic range* to accommodate quiet and loud speakers in quiet and loud environments. This problem is made more challenging by that previously-mentioned Lombard effect – whether they need to or not, people naturally speak more loudly in a noisy environment. So a naturally loud speaker working in a noisy environment is likely to speak extremely loud! The microphone must be able to pass this very loud speech to the recognizer without any distortion from overloading.

WHAT ELSE?

Let's assume we've built the world's best warehouse recognizer, making all the previously-described right design decisions. There's a lot more to do before we have an effective voice system for use in warehouses. Here are some other important criteria.

Adaptive Recognition

Thus far, we've assumed that users will speak the same way each and every time they talk to the recognizer. But, in practice, that isn't so. After eight hours of work, a user may be tired, or perhaps he came in with a cold. Such variations can change how a person pronounces words and, in turn, how the system interprets his sound.

An *adaptive recognizer* has the ability to continue to learn more details of the user's speech patterns as he or she interacts with it over time. For example, whenever the person speaks a particular digit, the system registers the response and fine-tunes its model of how the user is saying that word. Adaptation algorithms must be designed very carefully. Not all user utterances are good examples – we all have times when words just don't come out right, and in an industrial environment, some utterances will be corrupted by extreme background noise. The adaptation algorithm must take into consideration these anomalies and ignore them.
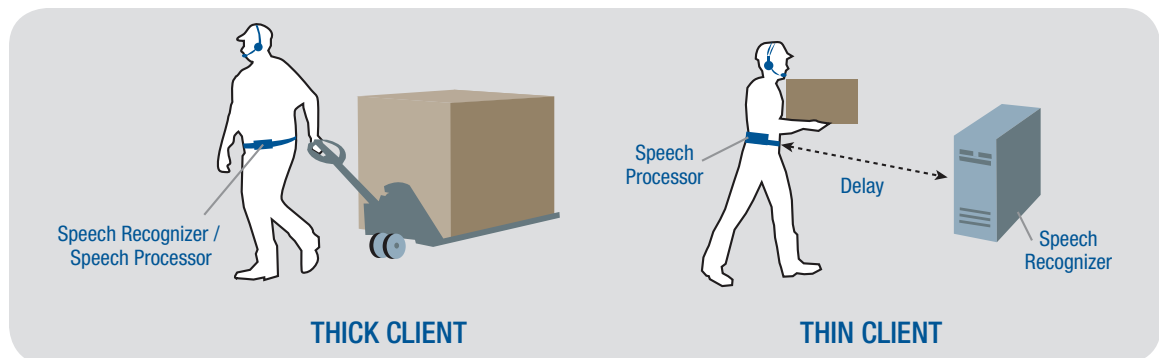
In comparison, an over-the-phone recognizer cannot generally take advantage of adaptive recognition, because it normally assumes it will hear a user only once. Since the next caller will have a completely different voice, there is no opportunity for the system to learn or adapt to a unique voice. Some dictation software programs do have a limited adaptive recognition capability that improves accuracy over time. However, adaptation is inherently less effective for a phoneme-based recognizer than for a word-based one, because in a word model, each component phoneme is adjusted to exactly how it is pronounced in that specific word, not a blend of how the phoneme is pronounced against all words. Adaptation at the phoneme level won't help if a user pronounces a word in an unusual way.

Responsiveness

Those old enough to remember making transatlantic phone calls via satellite will recall that the transmission delays, although quite short, made conversations surprisingly awkward. We humans are very sensitive to "conversational delays" when we don't have any visual cues to show us that

the other party heard us. Few things do more to reduce the productivity and overall satisfaction of a speech recognition user than a system with a slow or inconsistent response time. Increases of even fractions of a second to response times can make a pleasant conversation painful. If the user begins to think the recognizer hasn't heard her, she may begin to repeat herself. The recognizer, having heard her but been slow to respond, may interpret the new speech as another data item. Now the user and the recognizer are "out of sync." Delays and frustrations are inevitable.

For this reason, it is better for speech processing to be conducted directly on the wearable or portable computer that is with the user at all times. While it requires that these portable devices have significantly greater processing power, the gains in speed and performance provide this architecture, known as *thick client*, with a superior advantage. With thick client, the voice recognizer and the text-to-speech engine reside on the terminal. With each terminal's processing power fully dedicated to its one user and no requirement to move the speech data across a wireless network, a thick-client system can guarantee consistent and rapid response to its user's speech.

Speech Recognizer /
Speech Processor

THICK CLIENT

Speech
Processor

Delay

Speech
Recognizer

THIN CLIENT

The opposite of thick client is known as *thin-client*. With thin-client, the voice recognizer does not reside on the device. Instead, data must be transmitted to a server for processing and speech recognition.

This can cause delays, depending on network design, server software implementation, and how many users are attempting to interface with the server at the same time. As we've noted, even very brief delays will seriously degrade the user's experience with the product.

Thin-client devices are also completely undone by connectivity problems. Areas of weak Wi-Fi coverage in a facility dramatically increase the likelihood that data cannot be exchanged properly. Since a thin client requires nearly continuous high-speed communication between the device and the server, workers may have to return to an area of the warehouse where they can receive a stronger signal before proceeding with their assignments. Thin client is also more taxing on an IT infrastructure. Each device with marginal connectivity to the network can affect all of the other devices, by increasing the number of data collisions and retries. This can lead to productivity-sapping delays.
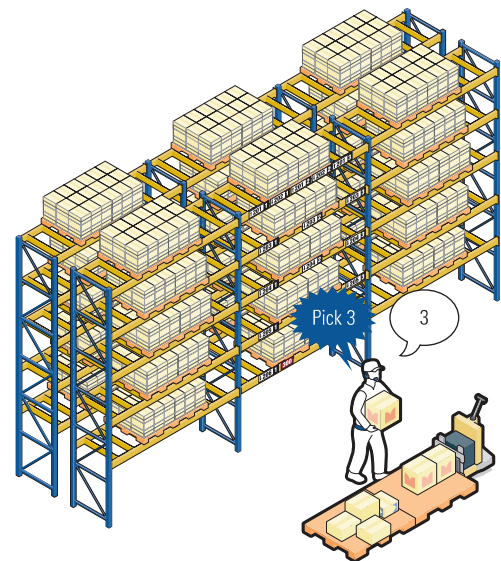
PUTTING IT ALL TOGETHER

Now that we have detailed many aspects of how the various speech recognizers work, let's further examine how they are applied in a warehouse workflow. It's important to note that even a recognizer that can distinguish one word from another with unerring accuracy and blinding speed may not provide a satisfactory experience for the user, if not designed and applied properly. Let's see how a warehouse voice system leads its user through a typical transaction, then we'll see what we have to build "around the recognizer" to create a high-quality user experience.

The *application engine* manages the prompts delivered to the user and specifies the user responses the recognizer should expect. It has access to information about what the user is to do, and the sequence of commands and responses that should be exchanged with the user. The application engine communicates with the recognizer via the *dialog engine*.

Let's take the common voice task of picking a warehouse order as an example.

Selection is very much like grocery shopping. Items are stored in *slots* in the warehouse racking. The wearable computer being used by the order picker (or *selector*) has a list of slots and item quantities that make up an *assignment*. The only real difference from grocery shopping is that the items to be picked are specified by location, e.g., "*aisle 3, slot 124,*" not by description, e.g., "*banana*."

After the selector completes a pick, the application engine looks in the assignment list to find the next slot to be visited. If it's in the same aisle as the current slot, the application builds a prompt for the selector that simply says, "*Slot 135*." (Warehouse voice applications are typically extremely concise – we don't want to waste the workers' time listening to unnecessary speech). The application engine sends the prompt to the dialog engine – an instruction for the recognizer to prepare to hear a sequence of (typically) three digits, and instruction that this prompt is not interruptible. The user must then listen to all of it before responding. A system's ability to offer both interruptible and non-interruptible prompts is essential. We must ensure that the user hears all essential information. However, at other times we may want the user to be able to interrupt the dialogue. For example, a user may ask for some kind of status report and then decides he doesn't need to hear all of it. It could be very frustrating if he can't interrupt and tell the system to move on.

Once the system issues the "*Slot 135*" instruction, the speech recognizer is started so that the system can listen to the response given by the user. In this example, the user would read off a *check digit sequence* to confirm the location. A check digit sequence is typically a two- or three-digit number physically attached to the facing of each storage position. Upon arrival, the worker must read off the correct number to confirm that he is at the correct location. The user reads the check digit, "*373*," and the recognizer interprets the spoken numbers and compares it to the models of the words in the active vocabulary that the system expects the user to say. If the recognizer finds a word or sequence of words it is confident in, then it passes the match to the application engine for processing. If the recognized speech does not match the correct check digit, then the prompt is repeated. If the recognizer hears a sound that it cannot find in the grammar, then it ignores the sound, assuming that it is extraneous noise that has entered the user's microphone (or perhaps the user talking to a friend about last night's game).

Warehouse voice systems must be "always on" and listening to the worker in case he needs to communicate with the system, perhaps asking for a prompt to be repeated, seeking a product description, or logging off of the system. Thus, the user does not have to push a talk button to speak (as is required on some voice systems). A button takes an extra step to push and does not keep both hands free. Because the software is continuously monitoring the user, the system must be carefully engineered to evaluate the noises it hears and to disregard extraneous noises or "off-topic" user speech, rather than interpreting them as intended responses.

For both maximum productivity and user satisfaction, the system should also minimize the number of words the user has to say. Some less effective voice systems require *anchor words*. These are words the user must provide to explain what will follow. For instance, before giving a check digit, the user may have to say "*Check*" to alert the system that the response following it will consist of numbers corresponding to check digits. Anchor words may also be added to the end of a sentence, such as "*Stop*" to tell the system that the response is completed. Additionally, a less effective voice system may have to add extra steps in its prompts to gain accuracy, such as confirming that what it thinks it heard it actually did hear, e.g. "*You said three, are you sure?*".

Let's continue with our processing example. Following confirmation that the user is at the correct location, the application engine gathers the next needed command, such as "*Pick,*" and adds the quantity to be selected from information relayed by the warehouse management system. The dialog continues as the user hears: "*Pick three.*" He confirms the quantity selected by saying "*Three,*" and again the recognizer compares the response to the word models and finds the best match. All of this complex process happens as quickly as the user can speak and listen, with no delays.

Once the user confirms the correct location and the pick quantity, the process repeats itself for subsequent selections. With a high-quality recognizer using adaptive technology to maintain its performance over time, and with a well-designed system surrounding it, task performance can be extremely high. Workers never have to slow or interrupt their primary function in order to interact with the computer.

ENDLESS POSSIBILITIES

In this chapter we have attempted to draw a solid line of separation between consumer and industrial speech recognition systems. As conveyed, the recognition accuracy of speech recognition technology in warehouse settings is light years beyond that of commonly experienced consumer applications. We have looked at the components and variations of warehouse speech recognizers, and we have examined a host of factors that directly impact the quality of speech recognition. Finally, we have made a strong business case about the value that can be assigned to maximizing a voice system's speech recognition quality. In today's world where companies must eliminate as much waste from their supply chain as possible, getting the greatest payback from a voice system via good speech recognition is a bottom-line issue and opportunity.

In our next chapter, "Solving Warehouse Challenges Using Voice," we will explore how voice technology has helped many companies solve common strategic business challenges. We'll also look at how companies using voice have experienced significant boosts in productivity, accuracy and safety, reduced training time, and improved customer satisfaction.

CHAPTER 3

Solving Warehouse Challenges Using Voice

In Chapter 1, we saw how voice is impacting distribution by increasing accuracy and productivity while reducing costs. We also explored the developmental history of voice from its earliest industrial applications to its modern widespread use in supply chain applications.

In Chapter 2, we demystified voice. We examined the different types of speech recognition systems that vary widely in their abilities to process information and provide desired accuracy. We noted how speech recognition technology used in distribution-based voice systems is completely different from the typical systems familiar to most consumers. We saw how being able to train a system greatly increases recognition accuracy and overall performance. And we also discussed how the various pieces of the voice puzzle work together to create a system that is highly accurate in its ability to recognize a user's speech and then leverage those voice responses for a productive workflow.

In Chapter 3, we will present a wide variety of companies from around the world that have used voice to solve their key business issues, and the many gains they have experienced with deploying a voice solution. Companies today face a myriad of challenges, such as an ever-changing business environment, SKU proliferation, and increased governmental and industry regulations. While productivity, or throughput, and accuracy are without a doubt the most important paybacks these companies have attained from voice, you'll also see how these organizations have reaped many other benefits such as increased safety, reduced training time, streamlined processes and improved efficiency through voice.

Finally, you'll get an inkling of how companies are expanding from using voice exclusively for picking to adding other warehouse functions, such as cycle-counting, replenishment and put-away. We'll cover these applications in greater depth in Chapter Four.

> "The block of granite which was an obstacle in the pathway of the weak becomes a stepping stone in the pathway of the strong."
> — Thomas Carlyle, 19th century writer and philosopher

FACING GIANTS

Every business faces obstacles. It is how a leader deals with challenges that determine success. Distribution today must be lean, productive, and accurate. If a company fails to gain efficiencies, it will not be able to maintain its competitive advantage and the leader will not maintain his or her position for long.

So, what are challenges that rob the sleep of supply chain leaders?

For many, they relate to a changing business environment, sometimes planned and sometimes reactionary. Just as the economy goes through ups and downs, with some doing better than others at any given time, distribution curves tend to reflect similar patterns. Some businesses experience tremendous growth, with the need to quickly add distribution capacity. Other companies are in more mature industries. For them, the challenge is to squeeze more from thinning margins. And still others face slowing growth or even contraction, which forces them to consolidate distribution networks and make remaining facilities more efficient than ever.

Voice systems provide tremendous flexibility to help companies adapt to a variety of business conditions. Due to their portable design, they can be used just about anywhere within a facility, as well at any other facilities that may be added or deleted from a network. This provides consistency that is scalable as conditions change. And the sweet spot for voice is in delivering gains in productivity and accuracy, which helps to keep costs low to deal with shrinking margins.

Good labor practices are also essential for success. Labor is the largest single expense in most distribution centers, so managing it properly can bring substantial revenue to the bottom line. As a result, labor management is where the most significant administrative opportunities lie. A voice system, working in conjunction with a good warehouse management system (WMS), can help companies manage their labor. The systems provide the data and the performance metrics to keep workers busy and productive, to weed out poor performers, and to provide incentives for exceeding objectives. Workers can easily move to where they are needed most, as they are not strapped to a fixed-in-place technology. This allows managers to make certain that workers keep working and products keep flowing.

Workforce aging is a further important consideration that managers in certain countries must face over the long term. Baby Boomers throughout the developed world, particularly in the United States, Europe and Japan, will not be able to retire as early as generations have before them. Distribution centers in these countries will have to adapt to older workers who may not have the physical capabilities found in younger workers. This will require warehouse designs to be more user-friendly. Voice can help here too, as its hands-free design allows older workers to more easily handle the cartons they must move.

SKU proliferation is another factor that adds to management headaches. Companies that once had to handle thousands of stock-keeping units today must deal with tens of thousands of SKUs. That means the handling and picking systems, not to mention the software that drives them, must be robust enough to track all of those products. It also means that storage and picking systems must be designed to handle the additional complexity that more SKUs require. Voice systems are inherently SKU-independent. They work just as well with 50,000 SKUs as with 100. And of course, the more the SKUs, the greater the chance for errors. Adopting voice helps companies manage the complexities a high number of SKUs brings, while still achieving exceptional service levels.

If a greater number of SKUs were not difficult enough, an additional problem is the fact that these SKUs are continuously changing. Seasonal businesses are accustomed to working with products with short shelf life, but a continuous churn of new products is prevalent in just about every vertical market today. The systems that handle and manage an ever-changing product mix have to be flexible, scalable, and adaptable, which voice clearly achieves.

Another challenge is the effect of increased regulations. All around the world, governments are increasingly demanding that companies distributing foods, beverage, pharmaceuticals, and certain health and beauty products be able to track and trace product lot numbers from the manufacturing step through the supply chain to the retail outlet. A chain of custody must be created for each order to meet newer regulations designed to protect consumers from tainted products. In the event of an emergency, consumer alerts and recalls can more easily be put into effect to thwart any potential public health threats. The need to know where any product is at any given time is a challenge that adds tremendous complexity to today's supply chains.

Along with track and trace requirements, the collection of accurate data is also needed for enhanced security measures in today's perilous times. Importers and exporters must now file additional customs statements detailing the origins of products and where those products have been throughout the supply chain.

While the exchange of data has increasingly been mandated by government agencies, customer demand for information is also increasing. It used to be that a customer would send a purchase order and not see any information on that order until he opens a carton and reads the packing slip. That is not enough today, as customers demand real-time visibility allowing them to see the current status of each product in an order, as well as the order history.

Voice's capability to capture and deliver data in real time allows for accurate and immediate information management. Advances such as the 'Voice Pick Code' endorsed in the U.S. and Canada by the Produce Traceability Initiative allow for the tracking of goods with negligible productivity impact. Collected data can be easily exported to document-reporting software that meets governmental requirements as well as customers' needs.

In addition to these demands for information, customers also want their orders delivered faster, with the right product, in smaller but exact quantities, on time, and in perfect condition. Voice helps users to meet these challenges by picking orders accurately and completely.

While the challenges that keep distribution leaders up at night are many, voice — by no means a panacea — is one solution that can not only help them survive, but excel in turning the obstacles in their paths into the stepping stones that lead to success.

Now let's look at some examples of how companies around the world have faced many of these challenges and found voice to be a major contributor in their supply chain achievements.

PRODUCTIVITY IS JUST THE BEGINNING

The primary goal of most warehouse upgrades is to make operations more efficient. But in today's distribution world, it is simply not enough to improve to the point where a company keeps pace with its competition. Systems have to show multiple benefits that allow performance to leap forward in meaningful ways. Wherever possible, they must be extensible — able to expand or adapt to all-new applications for the same investment.

Voice-directed systems have the capability to drive performance in a number of measurable areas. While many users have specific business challenges that they intend to address with their initial decisions to move to voice-directed distribution, they find that the systems improve processes in many unforeseen areas as well.

Companies almost invariably implement voice because of the huge gains it can bring in productivity and accuracy. It is not unusual for companies to increase their productivity by anywhere from 10 percent to even 100 percent in moving from paper or RF to voice. Similarly, accuracy can also rise substantially, with error rates decreasing by as much as a factor of five. Fewer errors lessen returns, improve customer service and reduce costs.

Take Sony of Canada, Ltd. for instance. This Canadian arm of the international electronics firm wanted to maintain its competitiveness by making its distribution operations more efficient to decrease costs and improve service levels. The company operates two facilities in Toronto and Vancouver for distributing Sony's wide range of products throughout Canada, including its consumer electronics and its line of broadcast equipment.

Sony of Canada initially used paper-based systems in its distribution operations, and then it began implementing RF-based systems before eventually adopting voice. Sony of Canada was actually the first Sony operation worldwide to implement voice.

"We wanted to have efficiencies through voice-directed work, such as being hands-free, which is a huge benefit for us," says Gary Davis, Operations Support Supervisor for Sony of Canada. "Voice makes the process very fluid. It's not a stop-start type of scenario like we have with RF, so that was very beneficial to us."

Sony of Canada pulls product from its warehouse as a batch, which it then sends to an area of the building where the batches are staged and sorted. Voice then directs "cluster-picking" of items to specific orders accumulated within the sorting area.

Since moving to voice, Sony of Canada has seen a 20 percent increase in throughput and it has also reduced errors by 20 percent. These achievements have allowed the company to reduce the number of temporary workers it adds during peak seasons. The operations have also netted better customer service, including an ability to address the order trend of customers wanting fewer quantities in each order, but placing those orders more often. This trend reduces carrying costs for customers, but requires that Sony of Canada be flexible to respond quickly when those customers need additional product.

"There is always the challenge to bring product forward, to get it picked, and to get it out quickly. We need to have a very efficient process to do that," says Rick Courtin, Supply Chain Business Process Manager of Sony of Canada.

"We rolled out voice in order-picking and cycle-counting and we want to get it utilized in all aspects of the operation," adds Paul Kirkpatrick, Vice President of Supply Chain of Sony of Canada. "The ability to do more with less at a reasonable cost to the organization was a simple concept that management could embrace," he says.

Japan-based IHI Construction Machinery Limited is another company that has made huge jumps in productivity since moving to voice-directed distribution. IHI is a leading manufacturer of large-scale construction equipment, such as excavators, hydraulic shovels, and cranes. Its Yokohama distribution center provides some 60,000 different replacement parts and supplies for its equipment.

IHI sought to standardize its work processes, simplify its operations, and gain efficiencies. Prior to implementing voice in June of 2008, the company used a combination of paper and RF-directed systems. But the RF system transferred its data as a batch, so there was no real-time information on current work. The paper, of course, also did not offer true visibility into operations. Implementing voice changed all of that.

"Real-time control of work processes has been enabled through Vocollect Voice without having to replace our existing ERP system," says Satoru Hanazono, Manager of the Information Systems Group at the IHI corporate administration office. Since moving to voice, IHI has standardized its processes and gained an average productivity improvement of 46 percent. Over a period of months, the company has also applied voice successfully for cycle-counting, receiving inspection, storage, and shipping inspection.

Along with significant productivity and labor gains, accuracy has also improved tremendously. After the first year on the system the company had cut work errors by 70 percent and had attained operating accuracy rates of 99.993 percent.

"The effect of these improvements in work efficiencies from voice is so great that I anticipate additional future benefits," adds Satoru.

Farmacias del Ahorro, Mexico's largest pharmacy chain, has made significant distribution gains. It operates six distribution centers and six cross-docking facilities spread throughout Mexico to supply 1,000 stores that cover 75 percent of the country. Selecting some 16,000 SKUs of medicines, health and beauty products and other merchandise was not an easy task. The company was experiencing a high number of mistakes in its selection process, suffered from a high staff turnover, and found it very difficult to control the inventories within its facilities.

To ensure confidence in what was sent out the door, distribution centers implemented a series of costly auditing steps within its processes. As a result, order fulfillment became a long and labor-intensive process.

The company determined that it could solve these problems with an integrated solution that included a warehouse management system, voice-directed picking, improved training, and the implementation of an employee incentive program to reward performance. It installed the voice system at all six of its distribution centers, with 180 voice-directed selectors on the overnight shift picking cases and split-case items.

The results of the initiative have been very impressive. Picking productivity increased 130 percent. Errors decreased by 85 percent and overall selection costs were shaved in half. Turnover has also been reduced dramatically.

KNOWING WHAT IT IS, WHERE IT IS, AND WHERE TO TAKE IT

A key to accurate order filling is accurate inventory. If a company does not maintain control of its inventory properly, then there is little chance that it will be able to meet customer demands of delivering the right product at the right time. Voice systems, working with warehouse management software and enterprise systems, help companies maintain users' inventories throughout the various steps of order fulfillment.

Cefrinor, a company specializing in cold product storage logistics, was facing major problems with productivity and accuracy in its picking process. Its distribution center in Salvador-Bahia, a coastal city located near the midpoint of Brazil's Atlantic shore, provides frozen and refrigerated storage for about 30 global and Brazilian customers.

The cold environment made it difficult for employees to pick from lists of paper. Workers not accustomed to the cold temperatures of the freezers, especially with the tropical air outside, tended to tire easily. This was complicated by the fact that many of the company's product packages and labels look very similar, making it difficult to quickly distinguish one product from another. These factors led to high error rates and lost productivity.

"Our workers made a lot of errors and gave us a very low level of accuracy," explains Luis Martinez, General Director of Cefrinor. "Fatigue was a factor. Reading (from paper) inside the cold storage, where we don't have the best lighting, was also difficult."

At first the company attempted to convert its paper-based warehouse management system to a new system that supported RF order selection. While this improved accuracy, productivity still lagged far behind. The RF systems also proved to be difficult for workers to use in the freezer environment and the units required constant maintenance. Cefrinor then turned to voice as a solution to solve both its accuracy and productivity challenges. Productivity soon rose by 30 percent and accuracy reached a level of excellence of 99.98 percent.

"The increasing level of productivity is proving that we made the correct choice with voice-directed distribution and we are very happy with the results. It's the productivity that pays for the investment," explains Martinez. "Cefrinor was the pioneer in Brazil in using a voice system in refrigerated environments. Today the company is planning to integrate voice with other workflows such as inventory management and quality audits."

For London Drugs, accurate inventory and order filling are a necessity. As one of Canada's largest retailers, the company has grown from a single drug store to a variety retailer with 75 stores in 35 cities throughout Western Canada. Stores sell everything from pharmaceuticals and over-the-counter medicines to refrigerators, electronics, and housewares.

Because medicines remain a major part of its business, accuracy has to be perfect at its two distribution centers. Without accuracy, the trust between its brand and customers would erode. "With the voice system, very few errors are occurring, and when they do happen, it's usually due to someone not following the simple procedures that have been put in place. And it's easy to correct," says Brian Best, Director of Transportation Warehousing and Distribution for London Drugs.

London Drugs had used paper and labels for picking its products, and with that technology incurred an error every 300 picks on average. Now voice directs picking of full cases and split case orders. While productivity and throughput have improved by 10 percent, picking errors have decreased to less than one per 1,000 picks with voice.

"I can't overstate the importance of a quick and accurate shipping system in variety retailing," explains Clint Mahlman, Senior Vice President of Retail Operations, Distribution and E-commerce at London Drugs. "With voice, our stores can maintain an in-stock position for customers while reducing the costs and headaches associated with picking errors."

North American grocery chain Safeway also faced issues of accuracy. Safeway has over 1,700 full service grocery stores that it operates in the United States and Canada, mostly on the west coast. The company has been using voice for its order selection for nearly four years.

"Our primary motivation was around selection quality," recalls Safeway's Evan Rainwater. "We already had engineered labor standards, so we did not expect as much in labor savings, though we did get a boost in that also."

Before implementing voice, Rainwater says that picking from paper yielded selection accuracies averaging about 98.2 percent. "When we got to over 10,000 cases and more, it represented a significant number of wrong picks going out of our facilities."

Safeway operates 15 full line distribution centers in the U.S., and now voice directs picking at all of them. Accuracy has since jumped to an impressive 99.89 percent.

"Voice has made a significant positive impact, as errors have gone down," adds Rainwater. "Our retail customers noticed a difference immediately and we had not even told them we were changing systems. Our claims reduced by 75 percent."

PROCESS STANDARDIZATION

Distribution inefficiencies often arise from the use of disparate workflows. One process may be utilized in one facility, while a completely separate technology may require work to be conducted quite differently in another. This especially happens to growing companies that add existing distribution facilities to their network as a result of an acquisition. Working with multiple systems can be very difficult, especially when many of the distribution functions are coordinated from a centralized office. Not only does this require the home office to process information from many different sources, but the lack of a common system does not allow for true comparison of performance from one facility to another and makes it difficult to set benchmarks that can drive performance improvements.

That is why many companies look to standardize their processes across all of their facilities. Among them is KPSS Inc., headquartered in Maryland in the United States and part of the German company KPSS GmbH, which in turn is owned by Japan-based Kao Corporation — a US$13 billion consumer products company. KPSS provides professional salon hair care products under the Goldwell and KMS brand names. The company has facilities worldwide and wanted to place the same technology in each to make them easier to manage.

"We have a very complex supply chain with many small warehouses. We required a solution that would help us to standardize and streamline our workflow across all of our distribution sites and countries," says Brian Hatfield, Vice President of Supply Chain North America at KPSS.

KPSS chose a voice-directed system operating on rugged mobile computers for picking and replenishment functions in its facilities worldwide. Voice is ideal for such workflow standardization, as it is scalable to nearly any size of facility. When combined with a standard warehouse management package, operations can be duplicated easily just about anywhere.

The first rollouts were deployed in distribution centers in Fresno, California and Toronto, Ontario, Canada with other sites in the United States, the United Kingdom, Finland, and Australia next to integrate the technology.

So far, results have been impressive. KPSS reports it has a similar process in all of its distribution centers. Voice has also helped to improve order accuracy by more than 75 percent and picking and replenishment labor has been reduced by more than 20 percent. The company was able to recapture its investment in less than a year.

CHANGING BUSINESS ENVIRONMENTS

Change, in business as in life, is inevitable. Or as the writer John Simone once put it, "If you're in a bad situation, don't worry, it'll change. If you're in a good situation, don't worry, it'll change."

There is no avoiding a changing business environment. It is how one copes with change that is the hallmark of a company. Very few businesses have not faced major hurdles within the past few years. A changing economy dictates that businesses alter course or suffer severe consequences. As a result, few businesses mirror what they were even just a few years ago.

Some companies have experienced rapid growth and expansion. Others have undergone business contraction, either selling off divisions or consolidating operations. And still others have had to drastically change their business models to stay competitive.

Fox Racing is a company that has experienced tremendous growth. This American-based company sells apparel and accessories to race fans, motocross and BMX riders, surfers, mountain bikers, and wake boarders. It also has a line of casual wear geared for the active lifestyle.

Because of explosive growth, including gains in its international business, the company's two California distribution centers were facing extreme orders that matched the extreme sports they served. Its paper-based systems could not keep up with demand, especially as orders were picked discretely, meaning each order required its own trip through the warehouse to complete.

"We had reached the point where we would be forced to turn business away if something didn't change, and change quickly," recalls Robby Dhesi, Vice President of Operations at Fox Racing. "We simply did not have the bandwidth to keep pace with our rapid growth."

Today voice is used for a number of applications at Fox, including cart-based picking, envelope-picking, replenishment, slotting, and cycle-counting. In the facility that does Fox's European distribution, voice is also used to direct receiving. And voice-directed packing is being tested in its U.S. operations.

Since moving these processes to voice, Fox's California DCs have doubled productivity, allowing the company to reduce the number of selectors from 35 to 18. Lines picked per hour have also doubled and accuracy is at 99.9999 percent. In 2010, the facilities recorded only 13 errors during the entire year. These gains have allowed Fox to experience a return on investment with its voice deployment of only six months, which was six months earlier than originally projected.

"Voice helped us reach a higher level of performance on every goal we set, and now we should be able to achieve our growth objectives across North America and internationally," adds Dhesi.

Smith Drug Company is another company with steady growth that has strained its distribution operations. This American wholesale distributor of pharmaceuticals and health and beauty products operates three distribution centers that serve more than 2,000 independent drug stores in the southeastern United States. The company's rapid and consistent growth required it to find ways to push more volume quickly through its buildings. But its growth in volumes was hampered by its reliance on paper-based processes. Paper proved to be slow and riddled with errors, not to mention the daily environmental impact of using reams of paper to print pick lists.

"It was obvious that we had to move fulfillment away from a paper system," recalls Randy McConnell, Director of Information Systems for Smith Drug Company. "That was holding us back, keeping us from reaching the kind of productivity and service goals we knew we were capable of. Goals we had to achieve to continue to be successful."

Errors in pharmaceuticals can be very costly. Some medicines retail for as much as US$1,000 and higher. It doesn't take many errors when handling such expensive products to eat into thin profit margins. Smith Drug tried a few remedies but found that errors were not eliminated, even with instituting a scan process.

Eventually, McConnell talked his upper management into trying voice. At first, Smith Drug maintained the paper system alongside the voice system. But that did not last for long, once the benefits of voice became apparent.

"I asked for a 20 percent increase in productivity with an accuracy rate of 99.99 percent and that was easily attainable," adds McConnell. "Now we are averaging about 80,000 units a day from our warehouses."

The installation of voice has allowed Smith Drug to absorb growth while improving its on-time deliveries. Shortly after installing voice, McConnell said that voice-directed distribution has brought his company an entirely new level of accuracy and accountability. "In the last month," he reported, "we had only one case that got on the wrong truck, and that's out of 7,000 cases a day. You can't beat that."

Diamond Comic Distributors went through a consolidation as a way of managing its growth. The U.S.-based Diamond is the largest distributor of English-speaking comics in the world. The company opened a 325,000 square-foot distribution center in Memphis, Tennessee in 2003 and quickly outgrew it within three years. It opened a second facility nearby, followed by a third building a year later. Its operations were slowed by the fact that it had three buildings trying to work together to do fulfillment.

"We had underestimated the space we would eventually need, and so we decided to bring all of our distribution under one roof," says Debby Salvatore, Director of Operations for Diamond Comic Distributors.

In 2009, Diamond consolidated distribution into a 600,000 square-foot building in nearby Olive Branch, Mississippi. Using voice for picking both cases and split case products, the new facility has provided Diamond with room to control its growth and the flexibility it needs to remain a leader in the comic book industry.

Order sizes vary greatly among the company's 25,000 SKUs, from one piece of a title to thousands. And since new editions are coming out weekly, the company needs the flexibility that voice offers to process orders efficiently.

"We had looked at different technology over the years, including RF and pick-to-light," explains Salvatore. "But we were not impressed with the flexibility these other systems offered. Our customer base is growing and each customer has different requirements. The others would not give us the flexibility to turn on a dime like voice can. Voice is a good fit for us."

Some 60 workers use voice on each of two shifts to perform picking. They work in two four-level pick modules where they pick titles by zone into cartons. Full cartons are also picked in the module and placed onto takeaway conveyors.

"Voice can handle just about anything," Salvatore says. "We have a lot of flexibility to start workers at any location within the pick module, in any zone, based on the order profiles."

Timeliness is extremely important in the comic book business. Each Wednesday new comics are presented on store shelves, and readers expect to see the latest titles in their local stores. That means each week's orders need to be processed efficiently to be delivered on time to customers.

"I can't imagine getting things out the door with the volumes we have without a good warehouse management system and voice. It is very, very critical," says Salvatore.

Another company that has been able to better manage a complex distribution network is The Co-operative Group, the UK's largest mutual (consumer-owned) business. It operates everything from funeral homes to insurance to food service, and is also the operator of the country's biggest convenience store network, with over 3,000 stores and 78,000 employees nationwide.

Handling the volume needed to supply all of the convenience stores was not an easy task. It involved many small depots that were difficult to service. Three separate distribution networks also existed — one for delivering ambient goods, another for temperature-controlled, and a third for frozen foods. Delivery zones overlapped, routes were inefficient or illogical, and many warehouses lacked sufficient space. The Co-operative Group realized that it needed to streamline its distribution network, ideally serving each store with a single, multi-temperature truck while improving the availability of goods in the stores.

"When you are dealing with small stores, it's a disaster if the product isn't available," says Trevor Ashworth, Director of Food Retail Supply Chain for The Co-operative Group. "You don't have the luxury of offering the customer a range of brands when you have limited shelf space — the product is either there or it's not. We knew that we were notching up a lot of unnecessary mileage with our deliveries and we had to find a way of fixing this problem while ensuring that product availability improved at the same time."

The Co-operative Group revamped its supply chain by building a new National Distribution Centre (NDC) for slow-moving ambient products. This instantly gave each of the regional ambient distribution centers more capacity. It also changed its transportation model to make deliveries more efficient.

To streamline operations within the NDC and the regional warehouses, the company installed a new warehouse management software system-wide and moved to voice-directed picking for filling store orders.

"We are well on the road to achieving our ideal network — multi-tiered, composite and strategically located," says Ashworth. "We have a framework that gives us visibility and control of stock and the flexibility we need to grow. The productivity gains have been very impressive. So far we have seen a 10 percent improvement in productivity at all the sites which have gone live with the system."

FLUCTUATING VOLUMES

In many companies, the amount of volume passing through their distribution centers varies widely. This is especially true for distributors with a high degree of seasonal business. Lawn and garden suppliers, candy makers, lumber retailers, swimwear manufacturers, and those who distribute holiday products are just a few affected by the need to push the majority of their annual volumes through their facilities in just a few months' time.

Signature Brands, part of Hero AG of Switzerland, manufactures Betty Crocker and Cake Mate dessert decorations and seasonal family decorating fun like Paas Easter egg decorating kits and Pumpkin Masters carving sets. These favorites of American consumers are often correlated with holiday seasons, which affect its distribution volumes.

"Because our business is seasonal, our shipping requirements vary tremendously from month to month, and nearly 99 percent of our orders must be shipped the same day they are received, so responsiveness and accuracy are critical," says Gary Stenzel, Senior Vice President of Operations for Signature Brands.

Prior to moving to voice, Signature Brands used paper for picking. According to Chief Information Officer Donna Dodson, "Since the voice deployment, we have been able to confidently handle our seasonal fluctuations. Labor assignments are more flexible, which has helped us to reduce our labor hours by over 20 percent, decrease inventory adjustments in picking areas by 80 percent, and cut our training time by 80 percent."

"Even before we had finished quantifying the cost savings obtained by deploying Vocollect and our new mobile computing terminals, we found that customer complaints had been greatly reduced," recalls Stenzel. "In my mind, that benefit alone is priceless."

MEETING REGULATORY DEMANDS

A major business challenge companies face today in most countries is meeting new government regulations. This is particularly a concern in the pharmaceutical and food industries where additional information must be captured on each item distributed. These regulations, also known as pedigree laws in the United States, are designed to protect consumers in the event of a recall, so that products can be removed immediately from the supply chain and the source of the problem can be quickly determined for resolution.

Working in conjunction with a user's WMS, voice can gather the data needed to conform to pedigree laws. At any point within the put-away and picking process, lot codes, expiration dates, and other pertinent information can be gathered and tracked by voice and the WMS as each product makes its way through the distribution facility. This information can then be shared easily as products move through the other parts of the supply chain.

One leading global pharmaceutical company launched an award-winning voice-directed order-picking solution from Vocollect within its highly regulated and validated pharmaceutical environment. The company has found that voice helped it track products with exceptional accuracy. It also reduced pick errors by 57 percent, which reduces the quantity of orders being returned. The voice system has also helped the company save 10 percent of its overtime costs and has reduced training time by 50 percent. It found that workers achieve maximum speed and productivity in four weeks rather than two months.

LABOR MANAGEMENT

Regardless of how much automation a distribution center has, it still requires people. People are the most important asset in any facility. Getting the right people and giving them the tools to be productive and happy in their work is an important piece to solving the distribution puzzle.

Labor remains the biggest single expense in warehouses, so it is essential that supply chain managers find systems that use labor more efficiently, particularly in picking areas, which are typically the most labor-intensive activities within the building.

Voice systems reduce labor as they drive productivity. They are extremely user-friendly, with operators becoming proficient much sooner than with other technologies.

VersaCold Logistics first introduced Vocollect Voice into Canada in the fall of 2009. The introduction was part of the implementation of its WMS, which is an RF-based system. "From our years of experience in the grocery business in the United States, we have proven there is a real ROI for the use of the voice technology on certain types of business. The decision to expand this technology into the Canadian warehouses was a no-brainer," said Robert Bascom, Vice President Operations, Eastern Canada, VersaCold Logistics. "One of the greatest benefits with using voice is the decreased timeline for new hires to get up to productive levels of work. Instead of focusing on accuracy, they can concentrate on safety, quality, and productivity."

"The speed and efficiency with which we can train such a wide variety of new employees to pick accurately and quickly, fresh off the street, is incredible. One person can train a number of new employees from start to finish in under two hours," said Henry van Oudenaren, General Manager of Ontario, Canada for VersaCold Logistics.

The company has since expanded voice to six of its 38 warehouses in Canada and has plans to add more facilities by the end of 2011. "The deployment of the technology is as easy as the deployment of RF technology," said Bascom. "Voice has helped to make us a world-class organization, and the level of accuracy we have achieved has made our customers the biggest benefactors."

VersaCold Logistics has many non-English speaking workers. Since Vocollect Voice is a speaker-dependent system (it compares spoken responses to words the user has previously recorded on a template), workers can speak back to the system with any language they choose, as long as their responses are consistent with their recorded template. Currently, instructions are given in English. Soon, the company will add the French Canadian language version for workers who prefer to receive instructions in their native tongue.

Language challenges also have not hindered workflow at Edeka Handelsgesellschaft Rhein-Ruhr, a German grocer that supplies 1,000 stores daily from three distribution centers.

"Our workers speak eight different languages in our distribution centers," says Denis Kleinhaus, Project Manager for Voice. "The training time, though, is very short because it is speaker-dependent."

Kleinhaus further explains that while instructions in Edeka's DCs are delivered in German, he has people who respond in a variety of different languages, including Turkish, Russian, and Arabic. "The system is so flexible that it is easy to add additional staff in case of future growth."

Even countries with comparatively low labor rates find that voice gives them competitive advantages through improved processes. In China, Qingdao Liqun Group is a large conglomerate with a diversity of businesses including transportation and third-party logistics. Liqun has been working to improve its quality of service and has moved to "smart" warehousing, including voice-directed distribution. Its facility in Jiaozhou is China's largest, most modern third-party facility with a campus that covers nearly 500 acres.

Today the facility operates completely paperless, with voice used for a variety of applications, including picking in cold storage as well as ambient areas, where it replaced paper and RF picking. Throughput in these areas since moving to voice has increased 28 percent and has achieved perfect orders. Training time for new workers has also decreased by 50 percent.

"The benefits we have seen from using the voice-picking system have been substantial," says Zhang Minggang, General Manager of Qingdao RuiTong Hi-Tech Co. Ltd., Liqun Group. "In comparison with RF picking, on average voice picking is 28 percent more efficient. This is far more than we had originally hoped for. To have achieved this level of efficiency in such a short space of time is very impressive."

That attention to work and not being easily distracted also solves another major business problem concerning labor — safety. When directed by voice, workers no longer have their eyes glued to a piece of paper or a screen. Their eyes are up and looking ahead at their work and at the area surrounding them as well. They are completely aware of their environment, including the lift trucks and other hazards that can cause injury. And since they are not looking down, they can continue to move while receiving instructions. They do not have to stop their walking while reading from a paper or screen. And while their eyes are free to focus on the task at hand, their hands are not occupied holding papers or scan devices. They are free to perform their jobs.

That benefit is extremely important to one French branch of a global company that distributes decorative paint products from three distribution centers in France situated at its manufacturing sites. Anyone who has ever carried a gallon or five-liter container of paint knows that it can be a heavy product to handle. This company moves some 36,000 pallets of paint each month, which adds up to about 20,000 tons. Handling that product hands-free is essential to safety.

For safety reasons, the company's operators must be equipped with ultra-minimalist equipment with the smallest, lightest form factor. Voice has been a perfect match for this organization's needs. Since using voice, the company has seen a decrease in warehouse accidents because pickers are eyes-free and hands-free, and they are more productive. There also has been a decrease in picking mistakes as well as greatly improved inventory control.

Not only are employees more focused on their environment with voice — they are also more focused on their work, which is directly related to the increases in productivity that voice users soon discover.

"One manager told me early on that voice really reduces the chatter in the warehouse, as workers are more focused on their tasks. While having a conversation with the system, a worker cannot be having a conversation with his buddy about last night's ballgame," says Ed Krupka, Chief Information Officer at Burris Logistics, the ninth largest cold storage warehouse company in the world, which uses voice for picking and replenishment. Voice has enabled 250 Burris users to help the company increase productivity by 25 percent and to achieve an accuracy of 99.97 percent.

"Voice technology really set us apart from other 3PLs," says Krupka. "We did not realize the full extent of the capability of voice until we began using it. We soon saw the flexibility it provides to reengineer our processes. We are able to respond much more quickly and are more agile."

Voice also gives Burris the data it needs to support engineered labor standards. "We have a defined metric for each time and motion, which becomes the standard for each task," Krupka elaborates. "Our employees are given incentives to measure up to that standard. Workers have confidence that they are tracked in real time and will receive the correct incentive pay for their performance. This allows our team members to perform to the best of their abilities."

One of the most difficult environments for using any technology is in freezers. Southeast Frozen Foods (SEFF) has successfully used voice in its freezer operations for many years. SEFF is both a wholesale grocer as well as a third party logistics warehousing and transportation company operating within 16 states in the southeastern United States. Its warehouses include freezer, refrigerated, and ambient storage. SEFF uses voice for a range of operations, including picking, replenishment, and put-away. Voice-directed truck loading is also currently being piloted at a SEFF facility in Virginia.

Working in the cold conditions of a freezer is among the most unpleasant tasks in distribution. It is very hard to work with paper labels, and writing on a piece of paper is nearly impossible. Scanners tend to freeze up or their windows become fogged, and it is extremely difficult to push the buttons on a scanner while wearing gloves.

As a paper-based operation prior to implementing voice, SEFF was experiencing shortages and inaccurate orders in its freezer selection.

"Our customers have an expectation that they will get the right product," explains Danny Payne, Vice President of Operations at SEFF. "We often had shorts because the labels would not stick well, or they would curl and sometimes stick instead to an adjoining box. When we printed the labels, they would be perforated at every eight labels. Sometimes we would find we were eight cases short because an entire section of labels would not get processed."

With voice, workers no longer have to cut the fingertips of their gloves in order to peel labels. Instead, they have both hands free to make the picks. The voice terminal is typically worn comfortably under their coats with just a wire leading to the well-fitting headset. The voice system performs flawlessly and is able to deal with the extreme temperatures, plus the noise-cancelling features filter out high levels of noise from the fans and compressors blowing in these arctic-like buildings. As a result, workers are comfortable and extremely productive.

Payne adds that language is also no longer a barrier. A large percentage of his workers speak Spanish, which made it difficult at times to understand English-written labels. Now workers have a choice of following voice instructions delivered to them in either English or Spanish.

"That makes the learning curve a lot easier," he says. "Our workers are very comfortable with voice and adjust easily to the technology. Looking at a label that had words in a different language was very intimidating to them. With voice technology they know that we care about them and want to make it easier for them to do their jobs."

EXTENDING THE VALUE OF VOICE

Vocollect co-founder Larry Sweeney states, "When we started applying voice technology in the warehouse, we had absolutely no idea it would come to address so many business challenges so well. We were just focusing on driving productivity and accuracy for our early customers. While that is still a central goal most companies have for the technology, all of our customers across so many industries and countries have taken voice to an entirely new level of value and performance."

In our next chapter, "Implementing the Voice-Centric Warehouse," we will take a look at how companies can reengineer their operations to take advantage of the many benefits voice can bring in workflow process improvements that extend far beyond selection, detailing the experiences and insights of customers across multiple industries and geographies.

CHAPTER 4

The Business Case for the Voice-Enabled Warehouse

In Chapter 1 of *The Talking Warehouse*, we discussed the history of voice-enabling technology – what it is, where it came from, how it developed, and when it became an indispensable tool for improving warehouse operations.

In Chapter 2, we delved a bit deeper. We saw in that chapter just how voice recognition technology works, and how some systems are better than others at delivering measurable results. We examined the advantages of trained systems versus untrained systems. Trained systems, which compare responses to recorded templates of the actual speakers, offer significantly more accurate performance.

In Chapter 3, we heard from the voices of success. We saw the many benefits of voice-directed distribution, including accuracy, productivity, and speed. We also examined how voice helps solve many common business challenges, including working in demanding environments, managing growth, designing flexibility into operations, ensuring a safer working environment, and providing ease of training. We also saw how many leading companies are putting voice to work in their particular operations to produce gains that have brought a measurable impact to their distribution activities, as well as a quick return on investment.

In this fourth installment of *The Talking Warehouse*, we explore how distribution operations can take voice to a step beyond its traditional role of picking orders. Next-generation voice systems are now being deployed as the centerpiece technology that directs operations in many areas of the warehouse. As voice enables the distribution center, the same high performance and cost savings achieved through picking with voice are realized facility-wide.

by Roger Byford and David Maloney

> "You have to be ready for a certain amount of process reengineering. Voice will cause you to rethink and redesign your workflows, and that is a good thing – because voice provides a real-time window on your operation like no other distribution technology currently available."
>
> — **Ed Krupka**
> **Chief Information Officer, Burris Logistics**

CUSTOMERS: THE CATALYST FOR CHANGE

The basic aims of distribution are still the same today as they were 100 years ago. The general process of placing the order, shipping the right product, in the right quantity, at the right time, and to the right destination hasn't changed. But the distribution facilities themselves have changed dramatically, especially over the past few years. And the major catalyst for this change is the need to meet constantly changing and expanding customer expectations.

Today distribution leaders face a perfect storm of pressures that require them to be extremely flexible and adaptable. Economic pressures and some improved efficiencies have reduced staff, so the workers left behind must work with tools that enable them to be ever-more productive. There is a continuing trend toward consolidating distribution operations, as borne out in a 2011 study that Peerless Media Research Group conducted on behalf of the Warehouse Education and Research Council. The study found that 33 percent of the nearly 600 survey respondents are consolidating and/or closing warehouses to reduce costs.

Moreover, supply chains now face increasing natural disasters such as earthquakes and tsunamis, political events such as strikes and protests, competitive pressures such as price wars and competitor changes, and ordinary operational problems due to lengthening supply chains, such as port delays. These supply chain disruptions can cause rapid changes in orders, requiring increasing flexibility both in staffing and in productivity required per hour. Couple this with increasing order volumes of small internet sales, expanding industry and government regulations and the ever-increasing demand for perfect orders, and it's easy to see why even continuous improvement of current processes and equipment – while important – isn't enough. Distribution leaders need to think of alternative processes and technologies that can give them immediate results and make the supply chain a strategic advantage for their businesses. In recent years, industry-leading distribution leaders have begun to realize the benefits of addressing these challenges using voice-enabled workflows across the entire facility.

BUILDING ON THE BENEFITS OF VOICE

When voice was introduced to distribution two decades ago, it found its first home in the order selection process. This was a natural fit, as order selection is the most labor-intensive activity in nearly every facility. Given the high labor costs, voice can quickly bring significant process improvements, with changes identified in a period of weeks and the investment recovered in months.
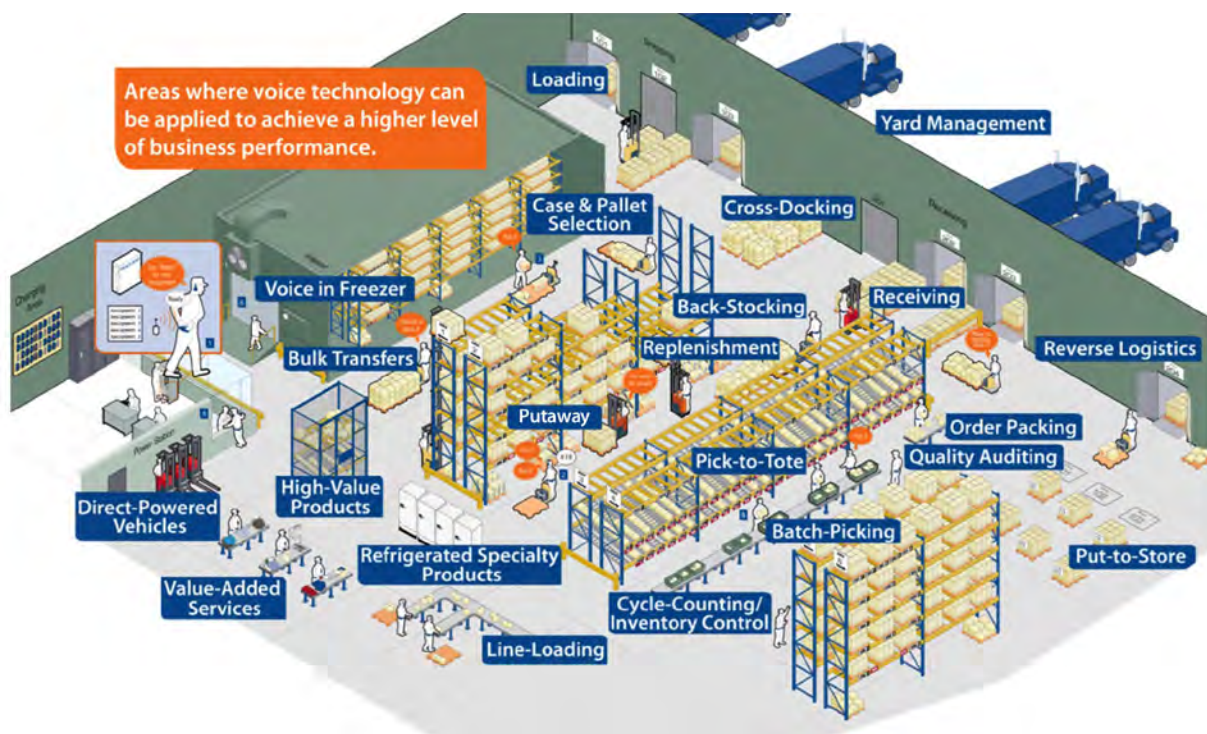
The eyes-free, hands-free benefit of voice has always made it ideal for the warehouse world, where all of the interaction with the system is through the headset and microphone through which workers hear instructions and respond to confirm actions – leaving their hands completely free to handle product. Their heads and eyes are up and aware of their surrounding environments, instead of looking at pieces of paper or screen-based devices. This improves overall safety and reduces potential injuries and workers' compensation claims. Having one's hands and eyes free is a tremendous advantage in receiving, putaway, shipping and many other areas where workers need to pick up and move cartons. Workers can accomplish these lifting and carrying tasks much more easily – and more accurately – if they are not holding RF scanning devices or papers in their hands or wearing RF scanners on their wrists and fingers.

Many companies rely on the real-time information voice provides. Managers enjoy having a factual view into each worker's individual productivity and task assignments. Having a snapshot of overall operations can be extremely important, as what happens in one area of the building can quickly impact others, often in a cascading manner. Order selectors cannot pick product if their slots have not been replenished. Packers cannot prepare products for shipment if there is a bottleneck in picking. Shipping cannot load trucks to leave the building on time if they are not receiving product from packing stations.

By having all areas delivering performance data in real time on a common platform, managers can respond as needed to bottlenecks, diverting workers where they are needed most to keep product flowing and efficiencies at optimal levels. A facility using voice as the centerpiece for all of these functional areas finds unmatched flexibility in managing its workforce, as employees trained on one common voice appliance can be moved quickly from picking to receiving, from putaway to replenishment, or from cycle-counting to loading. Whenever there is work to be done, the worker can easily slide to the new task by simply logging on into the newly assigned application. This makes it easy for companies to stay at peak performance throughout an entire shift. Once a worker becomes familiar with the operation of the voice interface and the operation itself, it is a simple transition to these new tasks.

Workers can use the voice system for receiving in the morning and replenishment in the afternoon. On the second shift, incoming workers can use those same devices to pick for several hours and then do truck-loading. When assigned a new task, workers do not have to leave their voice terminals behind in the picking zones in order to track down papers or handheld devices needed for receiving, packing or replenishment tasks. Instead, they can use the same voice device to work in any of these workflows. With a common platform and integrated software, the workflows are optimized, balanced, and coordinated at all times. And as we have discussed in earlier chapters, voice offers the easiest training process of any major warehouse technology. All workers have to do is listen, respond, and perform the task as directed.

As a company grows, the highly scalable voice system can grow with them, as depicted in the following illustration:

THE VOICE-ENABLED WAREHOUSE

COMMON INDICATORS THAT VOICE CAN IMPROVE BUSINESS PERFORMANCE

Workflows that can benefit from broad adoption of voice often exhibit common traits. Here are just some of the challenges that distribution supervisors, managers and industrial engineers can look for to see if a voice-enabled solution has merit for their business. If some of these issues do exist, voice may be able to improve productivity and accuracy, and in turn, address key business issues associated with customer satisfaction, business growth, SKU growth, and cost reductions.

- **High-cost, perishable or regulated merchandise –** Even a few shipping errors associated with high-cost, perishable or regulated goods can significantly affect profitability. For example, expensive luxury items such as designer clothes, accessories (purses, watches, etc.) and electronics; frozen or chilled meats, dairy, fruits, vegetables and even temperature-controlled drugs and medicines; or government-regulated items, such as pharmaceuticals, must be tightly tracked throughout the supply chain to avoid costly mistakes and errors.

- **High error rates –** High customer order error rates, customer claims and refunds, and internal error rates (for example, products in the wrong storage location, products in the wrong pick slot, or empty pick slots) indicate that the current system is not performing adequately. In today's environment, a business with high DC error rates won't succeed for long. The process evaluation and optimization that is part of deploying a voice-enabled workflow solution creates modified processes to improve accuracy, in a productive way.

- **Just-in-time shipments –** When the correct item must be shipped for immediate or just-in-time use, accuracy is paramount. For example, the service parts business for automobile and equipment manufacturers usually depends on the supply chain delivering the right part to repair a customer's automobile or a critical piece of production machinery. In other instances where just-in-time parts feed an assembly line, late or incorrect parts can bring manufacturing to a halt.

- **Idle forklifts or forklifts traveling without loads –** Idle forklifts are often a sign of inefficiency in the workflow in which the vehicle is being used. Vehicles traveling without loads, for example, on return trips, may indicate an opportunity for process improvement using voice. With voice-enabled vehicle workflows, voice sets the pace and vehicles are not idle or without a load unless they are being serviced.

- **Workers Who Move Between Workflows** – In smaller DCs, the same employees may perform receiving in the morning, picking during the day, and replenishment or loading late in the shift. Voice enables these DCs workers to use a single device for all of their warehouse operations, saving on capital equipment costs, maintenance costs, and training.

- **Overloaded or disorganized loading or receiving areas** – When the system is in balance, receiving and loading areas are kept clear. This is more challenging in receiving, due to the sometimes unpredictable timing of incoming shipments. Companies employing voice across multiple workflows have the flexibility to apply more labor in receiving when shipment volume increases. For loading, voice-enabled workflows can help to ensure that outgoing items are loaded onto trucks efficiently.

- **Work demanding heavy labor and/or dexterity** – Workflows that require employees to use both hands to push, pull or lift heavy objects lend themselves well to voice direction. For example, some DCs use heavy wheeled trolleys for receiving and loading. With paper or scanning, at least one hand is needed to hold the paper or scanner, leaving just one hand available for other tasks. With voice, both hands are free to do the "heavy lifting."

- **Processes are not interleaved** – In a DC that has the opportunity to improve operational efficiency through interleaving, voice provides an excellent solution. If the interleaving takes place within one shift (e.g., putaway and replenishment), voice can be used to direct the (typically vehicle-based) employee from one workflow to another. If the interleaving occurs from one shift to the next, voice can be applied to all workflows and the worker only needs to be trained on one system.

- **Shorts ("mark-outs") and bottlenecks or flat productivity improvement in picking** – When a stocking shelf is depleted ("short"), replenishment is not in sync with picking. Often this problem can be resolved by voice-enabling the replenishment workflow along with picking. Relieving shorts can improve picking productivity and reduce the need for "scratch runners" – last-minute pickers – to pick up missed product at the loading dock. Having multiple workflows on the same voice-enabled technology can allow supervisors to shift employees easily between workflows to relieve bottlenecks in picking.

- **Relatively short travel distances and/or dense picking –** Since voice does not improve travel times, per se, but can allow workers to talk/listen while traveling, the gain from applying voice in workflows one with relatively short travel distances will be higher than one with longer travel distances. A few leading-edge companies with relatively dense picking (usually case quantities with small cases) have achieved dramatic improvements in productivity in zone-picking by using voice to enable multiple workers to pick efficiently within the same zone.

- **Clustered and idle workers –** A DC with idle workers is typically indicative of one that is out of balance or is not taking advantage of the potential of interleaving processes. Implementing voice across the DC, together with process improvements, can bring the system back in balance and enable process interleaving. The result is better utilization of manpower and fewer idle workers.

- **High number of quality auditors –** Quality auditors are typically deployed in DCs with high error rates, with the objective of identifying the source of the errors. With a voice system in place, accuracy rates improve, reducing the need for dedicated quality auditors.

- **Workers spending time looking at paper, screens, and keyboards –** Fundamentally, voice improves productivity by freeing up a worker's hands and eyes. The amount of time a worker spends looking at paper and screens or manipulating a keyboard may seem negligible, but when these operations are performed hundreds of time each day, the lost productivity adds up quickly. Some studies have found that up to 20 percent of each worker's time is spent on these non-value-added activities. By removing these obstacles, voice lets workers focus on the primary task of moving goods through the DC and to the customer.

THE QUEST FOR MAXIMUM WORK PROCESS OPTIMIZATION

Installing a voice system, by itself, does not capture the full benefits of a voice-enabled warehouse. The biggest gains are realized when voice is coupled with process improvements. Voice-directed work enables process optimization that can reduce time per task, combine tasks between workflows to eliminate work altogether, and interleave tasks to eliminate unnecessary downtime between tasks. To optimize for greatest efficiencies, products may have to be reslotted. The layout of the storage and picking areas may need to be rearranged to improve the flow of product through the facility and to optimize space utilization. Product flows may also need to be adjusted to lessen travel times.

To achieve the greatest possible effectiveness, managers must look at every process within the building and be willing to reengineer them in ways that bring out the greatest gains from their people and equipment. With voice at the center, efficiencies and accurate handling can be achieved in all areas of the operation. All interactions between the computer that handles the bi-directional flow of information and the human who performs the work are optimized.

The chart below illustrates how voice has helped companies meet stringent performance demands in workflows beyond picking.

HOW CURRENT VOICE USERS MEET STRINGENT PERFORMANCE DEMANDS WITH VOICE

		Receiving	Putaway	Replenish	Put-To-Store	Outbound	Inventory/ Cycle-Count
Problems Solved By Voice*	For All Workflows	• Greater accountability (both accuracy & productivity) • Improved safety • Eliminating exceptions (especially multi-modal implementations) • Reduces overtime and enables temporary hire training during peak periods • Reduces bottlenecks through switching employees between workflows					
	Particular To Workflow	• Improves inventory tracking • Enables additional data capture • Less downtime between shifts	• Greater efficiency during peaks • Speed to shelf • Less downtime between shifts	• Easier cross-training between workflows • Reduced shorts of fastest-moving goods • Enables interleaving with put-away • Decreases wait time in picking	• Greater count accuracy • Easier volumetric pallet building • Enables additional data capture	• Reduced mis-ships and shorts • Enables consolidating pallets • Enables additional data capture • Improves meeting delivery timeframes	• Enables interleaving with other applications • Increases accuracy through hands-free counts • Eliminates costly and time-consuming periodic audits
Common Traits of Current Users	For All Workflows	• Interest in a disciplined and standardized approach • Desire for cross-training employees and interleaving workflows					
	Particular To Workflow	• High throughput • High-value products • Regulatory requirements	• Peak hours or seasonality • Spoilage or fragile products (frozen goods) • SKU turnover	• Fast-moving Consumer Goods and Grocery verticals • Cyclical day	• Already implemented profitably in selection • High downstream demand variability	• Stringent customer requirements • High cost of returns • Regulatory requirements	• SKU turnover • Stringent customer requirements • High-value products

SOURCE: Vocollect research

* Productivity improvements are stated versus paper unless otherwise noted

Let's take a look at how voice can be effectively deployed in common distribution workflows. We will start at the point where products enter the building – receiving.

RECEIVING

Receiving is where the arriving product is brought to be formally accepted and validated into the warehouse or distribution center. Goods enter the DC through a variety of receiving processes. Most receiving operations consist of trucks arriving at the DC from multiple suppliers at different times of the day. The goods from these trucks are unloaded into a holding area. Receiving workflows typically involve inspection, counting and other manual operations. Using voice enables these tasks to be performed efficiently, since the worker's hands and eyes are free to perform the tasks instead of being used to manipulate paper or scanners.

Some receiving operations require additional quality controls. Ideally, advanced ship notices should be sent by shippers to notify the facility of incoming products. If properly coordinated with the vendor, receipt-friendly labels will also have been placed on arriving pallets and/or cartons prior to shipment. Upon arrival at the facility, products are unloaded from trucks into the receiving area. Using voice, a worker can then read into the system a special code, such as a vendor ID or shipment code, which is located on or sometimes within the license plate number (LPN) on the pallet's label. Oftentimes workers use a scanner for this part of the receiving process. This is matched by the WMS to the advance ship notice, alerting the software that the shipment has arrived. Some DCs require workers to identify a particular pallet among the receipts. To accomplish this identification quickly, the voice system typically prompts the worker to read the last three digits of the pallet's license plate number. In the unusual case where there is more than one pallet with the same three-digit ending, the voice system asks the worker to select the correct one.

The voice system prompts the worker to confirm the number of cartons on the pallet and may direct the worker to open cartons and inspect certain items, place a facility label on each carton, or perform other tasks specific to the facility operations. If the incoming pallets are stacked too high for shelving, truck drivers or receiving dock workers may have to "lump" the product, or move cartons to another pallet or empty pallet in order to lower the pallet height to fit into storage. The voice system can help to manage the information about this process. Voice may direct the worker to set aside some pallets or cartons for further auditing, especially when dealing with newer vendors.

Additionally, some cartons may be directed to cross-dock operations. The receiving worker may be instructed to label these cartons with a shipping label, and place them onto either a pallet or a conveyor for transport directly to shipping.

Voice also offers the flexibility to gather additional information about the product as part of the receiving process, assisting operations further downstream. The lot codes of some receipts, for instance, may need to be recorded for track-and-trace purposes. This information is required in many parts of the world for products such as foods and pharmaceuticals. Voice may additionally prompt workers to measure and speak the temperatures of incoming products or refrigerated truck beds. This information is recorded into the management software to ensure that nothing has been spoiled or damaged during transit. Voice can also be used to record the dimensions and weights of receipts, with this information later used for slotting, cubing, weight verification during picking or loading, and invoicing purposes.

In some cases, it may be desirable to combine some scanning with the receiving operations, particularly when the identifying number of the carton has too many digits to speak easily into the system. In this case, a ring or wrist-mounted scanner can be used in conjunction with voice. The voice system simply prompts the worker to scan the barcode at the desired time. Today's voice systems are designed to integrate with a wide variety of scanning peripherals, many of which can save the DC operations money versus having to use a full-function mobile computer.

The captured data, whether gathered by voice or scanning, provides information on that receipt for all of the processes to follow. The information can also be used for the implementation of standards to help managers evaluate individual labor performance.

The August 2011 *DC Velocity* Magazine article, "Milking the System," profiled the experiences of Glanbia plc, an international nutritional solutions and cheese group company headquartered in Ireland.

> When Glanbia converted over to voice for its Irish dairy business, it didn't stop with picking. It found ways to use the technology in virtually every part of its distribution operation. What started Glanbia down this road was its growing frustration with the manual processes it was using at the time.

The company installed voice systems from Vocollect at four sites. The voice system now performs a number of operations in two milk manufacturing facilities and two distribution centers. One of the DCs distributes milk and cream products, which are bottle- and carton-based, while the other deals with other food products that are primarily case-based.

The central milk warehouse turns its stock three times each day, so it doesn't actually store product – it merely stages it so that it can be picked for delivery by smaller route trucks. The facility operates 24 hours a day, six days a week, turning out 1.6 million liters (approximately 422,675 gallons) of milk each day.

To ensure quick flow-through, voice is used in a number of processes within the central milk warehouse, including receiving products into the 23,000 square-foot facility. The "goods-in" receiving process starts with the driver's logging onto the voice system upon arrival. The voice system then prompts him to read off a delivery number listed on his dispatch paperwork. As the driver begins unloading the truck into the goods-in receiving area, he reads into the voice system the trolley's carrier number, which was attached to each trolley before it was shipped from the milk manufacturing center. This marries each trolley with the receipt.

PUTAWAY

In putaway, operators take pallets from a receiving location to a specific warehouse location. Once information is gathered on the receipt, the voice system receives putaway instructions from the warehouse management system. Typically, a different worker also wearing a voice device will be summoned to pick up the load. Since the putaway function usually requires a lift truck, DCs often employ truck-mounted voice terminals for this task. These terminals draw a small amount of power from the forklift battery, providing for extended voice terminal usage over a single battery. Many of these terminals use Bluetooth to wirelessly communicate between the device and the headset worn by the worker. That allows him to leave the vehicle and still be in contact with the system without having a tether to get in the way.

Products can be taken to several different destinations for putaway. They can go to reserve storage racks where they are held until needed, placed into automated storage, or taken directly to pick faces if the WMS determines that a slot that will soon be needed for filling orders is currently low in inventory. Voice will relay instructions for all of these putaway areas.

With voice in the putaway function, the driver typically receives instructions and provides confirmation while driving. This process eliminates the time required to manipulate a scanner and/or view a vehicle-mount unit screen. In addition to improving productivity, accuracy is improved, because the voice system will not allow the user to move to the next task until the goods have been put in the appropriate location.

When the driver arrives in the receiving area, he either reads off the identifying code on the pallet license plate or, if the load has been assigned to a staging area near the dock, he is given the location and reads off a check digit to confirm the pick-up location. The driver is then instructed to transport the load to the proper putaway destination.

The advantage of using voice for this function is that the driver is eyes-free throughout the entire process. He can be driving while he is hearing the location. He does not have to stop to look at a screen, saving valuable time and avoiding the hazard of looking at a screen while driving. Upon arrival, the driver reads the check digit off the location and puts away the load. In some self-directed designs, drivers themselves determine putaway locations, often choosing the closest available slot. In these instances, the driver is directed to read off the location number and check digit to inform the system of the slot he has chosen. In both cases, the check digit prevents lost pallets, which can be a huge cost issue, especially with perishable, seasonal or rapidly depreciating inventory. Inventory is immediately updated so that the product can be recorded as available for replenishment or readied for picking.

In the Glanbia example from the *DC Velocity* article,

> A goods-in person will then complete the putaway process, also using voice. The worker will read off the carrier number from the trolley (carts with shelving used to hold milk; milk is placed into the trolleys at manufacturing), and the voice system – together with the SAP warehouse management software – will direct him or her to take the trolley to an assigned staging location for picking. The worker can also choose his or her own location by informing the voice system of the change. A check digit posted above each location must also be read to confirm the load is put away into the correct area. The worker confirms the putaway, the status of the load carrier is updated, and the stock is then available for picking.

REPLENISHMENT

Replenishment is when operators restock a location, typically by moving product from a reserve storage location to the replenishment location. The replenishment process is often a combination of putaway and picking in the way the task flows. It is similar to picking, in that the worker is directed to a particular location to gather a product. However, instead of filling an order, the product is needed for another operation, such as refilling flow racks in the pick module, moving product to value-added services, or taking product for batch processing. Once it arrives at its destination, the product is handled similar to a putaway flow.

The replenishment process begins with the voice system directing the lift truck driver or pallet jack operator to the storage location where the needed pallet resides. Again, the driver can be moving while the voice system gives directions. In optimized designs, the WMS will assign the closest driver available to the storage rack. Once there, he reads off the check digit to confirm he is collecting product from the right location. The system will then direct him to gather a pallet, or if less than a pallet load, the needed number of cases. The driver confirms the selection and while he is pulling the product, the voice system tells him the destination for the load. The driver then takes the load to the location, reads off the check digit for the slot, and deposits the load.

Voice-enabling facilities for putaway and replenishment tasks pays big dividends. An operation can expect to see a 20 percent improvement over paper in staff productivity for these tasks while also increasing volumes and accuracy. There is also as much as a 65 percent and higher reduction in training time in using voice for these functions compared to other methods, and facilities can save greatly in capital expenditures by deploying voice terminals instead of other technologies. Many warehouses also see a dramatic improvement in productivity in the selection workflow. This improvement happens because the WMS directs replenishment workers to pick slots that are running out of inventory before the pickers arrive at those slots, reducing the amount of time the selectors spend either waiting for a replenishment or moving on to other picks and then traveling back to gather the shorted product once the slot has been replenished.

Headquartered in Monmouth Junction, New Jersey, Sysco Guest Supply is the largest full-service provider of personal care amenity programs and room accessories to the lodging industry, servicing virtually every large hotel worldwide. The company uses voice enablement in 10 distribution centers. Sysco Guest Supply is owned by Sysco Corporation, the global leader in selling, marketing and distributing of food products.

Says Director of Distribution Jim Makowski, "For a mid-size company, each of our regional distribution centers selects and ships extraordinary amounts of inventory. This volume generates a significant amount of replenishments. We needed a mechanism that would allow us to prioritize these replenishments. We developed and designed a process called 'Item demand.' Through our WMS and the voice system, we aggregate quantities by item to be selected and shipped that day, so that we can prioritize the replenishments in that manner. Additionally, we over-fill our forward pick locations (if necessary) to satisfy the demand for the day. This ensures that we are selecting everything from the home floor level locations, while requiring only one replenishment daily. The voice system indicates whether the worker is performing an item demand or general replenishment move as follows: "item demand – go to location number XXX" or "replenishment – go to location number XXX". The worker validates the location number check digit and product identifier, and is then given the quantity to select from the reserve location.

By using voice for replenishment, Sysco Guest Supply has realized a significant 80 percent efficiency gain.

Another strong proponent for voice in replenishment is Ed Krupka, Chief Information Officer of Burris Logistics, the ninth-largest third-party logistics company in the world.

"There are two strong benefits to using voice for replenishment that we have experienced. First, this gives distribution management and the individual worker real-time access both to make requests for replenishment and to get updates on the status of requested replenishments. Second, it enables a real-time reassessment or recalculation of overall task priorities in the warehouse. It also greatly reduces the loader's wait time, not having to wait around for so many stray out-of-stocks to come to the loading dock," Krupka says.

Another important advantage in using voice for both putaway and replenishment is the opportunity to do interleaving of tasks. Interleaving means that once one task is completed, the system may direct the worker to perform a different type of task within that same general area. This eliminates dead runs where the driver returns to the starting point with his forks empty.

Let's take an example of a putaway task that also involves interleaving. Once the driver deposits his load, the voice system can easily relay a replenishment order required from a nearby storage position. He pulls that load and takes it to the desired location before heading back to receiving to gather another load for putaway. In a similar manner, he can also pick full pallets needed for orders while doing replenishment and then take these loads directly to shipping. The idea is that drivers are not locked into specific tasks, such as only doing putaway or only doing replenishment.

Instead, using the common technology of voice, there is the opportunity and flexibility to easily move between tasks to where work is required, where travel time is reduced, and where the greatest productivity and efficiencies can be obtained. These efficiencies also apply to truck-mounted screen and keyboard terminal applications, also called vehicle-mounted units [VMUs], although VMU technology lacks the safety and eyes-free travel time advantages of voice.

Fox Racing, the U.S.-based manufacturer of apparel and accessories to race fans, motocross and BMX riders, surfers, mountain bikers, and wakeboarders, has interleaved its voice-enabled putaway and replenishment functions. Doing this, Fox saw a 20-25 percent increase in productivity, moving from putting away 25-27 lines per hour with RF scanning to approximately 35 lines per hour with voice.

PUT-TO-STORE

In put-to-store, workers pick one and more license plates (which identify what product and what quantity is in the container) and put product from that license plate in one or more locations.

Many retailers and other suppliers who distribute to their own stores or ship regularly to customer stores can use voice to direct put-to-store activities. In a way, this activity can be considered almost a combination of replenishment and batch-picking.

A source carton or pallet arrives into the put-to-store zone holding products to be divided into multiple destinations, or "put" locations. These locations are typically pallets, cartons or totes arranged in racking or at floor-level staging locations. The voice system directs the placement of a number of items from the source tote into the many destination (order) totes. Check digits at each order tote are used to confirm the destination, and the worker also confirms the number put to each location. Using voice for this application can be a very fast way of filling store orders. It provides a consistency of process and a high degree of accuracy. For DCs with a WMS that allows volumetric analysis, voice-based put-to-store simplifies the process of building pallets to the right dimensions for the truck – a particularly useful function if the truck heights or destination loading docks or stores have different pallet height requirements.

Tiendas Comercial Mexicana, S.A. de C.V. is the third largest supermarket chain in Mexico, with 200 stores in six different formats serving over 50 cities. The company operates four distribution centers, larger ones for dry goods and perishables and two regional facilities. In its put-to-store process, the picker takes the complete pallet from the receiving stage. The system sends pickers to each stop and the voice system specifies what quantity should be placed. The pickers dictate

the variable weight of each delivered box. When the pallet is full, they follow the same process of stretch-wrapping and shipping. If possible, they load the next product directly onto the truck.

Variations in weight are important; sometimes if goods or produce do not have a set weight, an average weight is used. With meat products, it is easy to weigh the meat and record the net weight with voice, whereas with RF scanning, reading and typing the weight caused too many errors and wasted too much time. When the put-to-store workflow was voice-enabled, the workers experienced a 50 percent-plus productivity improvement.

Says Carlos Ramos, Corporate Logistics Director at Tiendas Commercial Mexicana, "The ability to interleave tasks through voice helps us put more people on receiving or put-to-store during high-volume order periods that typically occur at our peak pre-summer season. This flexibility has helped us to avoid bottlenecks and reach a level of productivity that has allowed us to manage seasonal peaks only with overtime hours – a major savings over having to hire and train seasonal temporary labor."

OUTBOUND AND LOADING

This is generally the process of putting the product load that is now in the staging area onto the correct truck in the right order for delivery to multiple locations.

Once an order has been completed in picking, a worker is instructed through the voice system to take the tote, carton, or pallet to shipping. Sometimes this is done manually using carts or by lift truck, and at other times a conveyor handles transport.

Voice also can be used to sort orders for shipment. The voice system can instruct the worker where to place the product for staging before it is loaded onto outbound trucks. It can also direct the loading of products onto pallets in a particular stacking order that will ensure their strength and integrity during transit or, similar to put-to-store applications, vary delivery to a particular truck by pallet height according to transportation requirements.

It does no good to pick orders accurately, only to have them loaded onto the wrong truck or to have them left behind from their designated shipment. The cost of re-shipping pallets is considerably expensive, and sometimes not factored into the DC cost because it is included under transportation. The problem of pallets left behind is particularly common at facilities with more complicated loading docks, such as those with two different loading locations for refrigerated

versus ambient-temperature pallets. Voice also can help in this area, directing the loading of outbound vehicles. Instructions for loading are delivered through the worker's headset. A check digit on the dock door verifies that the load is placed on to the correct vehicle.

The voice system can also direct the loading in a particular sequence. For trucks making multiple stops, this sequence might be in the order of delivery so that the products needed first are loaded onto the truck last. Sequenced loading can also be used to balance loads, making certain that weight is evenly distributed throughout the trailer and the load is cubed to take advantage of all available space.

Facilities using voice for truck-loading report significant increases in staff productivity for this task. Because trucks can load quickly, doors turn faster, the need for floor staging space is reduced, and the facility is better able to cope with increasing volumes.

As outlined in the *DC Velocity* article on Glanbia,

> Voice also directs the loading process. When an order is ready to ship, a loading person logs onto the voice terminal. He then reads a delivery docket number found on the dispatch sheet (similar to a packing slip) that is later given to the driver and accompanies the order in transit. The worker is then directed to load the truck in reverse delivery sequence, according to the unit carrier label attached to each trolley. The worker must read back the last four digits of this number to confirm that the correct trolley is being loaded onto the truck in the proper sequence. The process continues until the entire truck is loaded.

CYCLE-COUNTING

Besides managing the main functional areas of the warehouse, such as receiving, storage, picking, and shipping, voice also can optimize other activities in the warehouse. Voice can support inventory control, also known as cycle-counting, the process of counting product at specific warehouse locations. This tedious process is made much faster, easier and more accurate using voice. With voice, the worker's hands and eyes are free to count product. They do not have to fumble with paper or a scanner while trying to determine the amount of product in the location.

The process for using voice for cycle-counting starts with the worker directed to a location. Upon arrival, she reads the check digit to confirm the location, and then reports to the system the amount of product in the slot. In some systems, cycle-counting can be interleaved with picking or replenishment tasks. The voice system can direct the worker to report on the amount of product left in the bin after a pick, or the same inventory count can be asked of the worker when she restocks the slot. With voice, as with screen-based terminals, the system can immediately ask for a recount if the level doesn't match. That can't happen with paper.

Sony of Canada, Ltd. performs cycle-counting with voice on a daily basis. Its WMS is set up based on an ABC type of methodology, with the fastest-moving foods assigned an "A" and slowest moving goods assigned a "C." Cycle-counts are generated and operations staff are sent instructions via the WMS to the voice system to perform the counts. Location and SKUs are scanned and counted as blind counts. If the count does not agree with the system, the operator is asked to verify again. If the second verification doesn't agree with the system count, an audit count is automatically generated. These audit counts are done by supervisory or management staff only.

According to Sony of Canada's Business Process Manager Rick Courtin, "Using voice for our cycle-counting is a much more focused and accurate way to perform the task. In addition, the hands-free feature of a voice system saves a tremendous amount of time."

Fox Racing has netted a 90 percent productivity improvement by using voice to interleave cycle-counting and picking. Fox's auditors require the company to have inventory counted every 90 days. According to Robby Dhesi, Vice President of Operations at Fox Racing, "Interleaving voice-picking with voice-cycle-counting picks up 30-35 percent of the cycle-counts during each 90-day period."

Fox turns on and off its cycle-count activity strictly based on order volume as calculated by its WMS. The company has three separate ways to pick up cycle-counts: 1) In picking, cycle-count never gets turned off based on a set quantity; if the SKU falls below three and a picker is directed to the location, she counts to zero. "We want to have complete accuracy in the warehouse, when if the WMS says we only have one left, then we truly only have a quantity of one. That way, if the picker needs one item, then we know it's there to pick," says Dhesi. 2) Fox still uses full RF scanning for cycle-counting for more scheduled cycle counts, turned on a time period setting of every 90 days. Every location is counted in full every 90 days, and this process picks up locations that haven't been counted in between picking activity. 3) Fox also has voice-enabled cycle-counting for ad-hoc checks of individual locations. Continues Dhesi, "Most people use voice for picking, which we did, too – but we found that the pickers were standing around waiting for the next order or cart to get there. So instead of them waiting, now we have a function to do a 90-day scheduled cycle-count that gets triggered if they are waiting for an order. All the picker has to do is read the aisle in which she is currently standing, and she is directed to cycle-count any location that has not been counted in 90 days." Now Fox Racing does less than 10 percent of scheduled cycle-counts in the normal 90-day cycle, because most locations have already been counted.

OTHER WAREHOUSE FUNCTIONS

Similar to cycle-counting, but typically performed by a separate group, is the auditing function. This is typically conducted by supervisors or team leaders, primarily for auditing counts at particular pick faces to verify the accuracy of pickers' counts, and/or auditing a subset of finished pallets or bins after picking to verify that the numbers and items picked were correct and the counts were accurate. Many voice customers have been able to eliminate periodic inventory audits because of the accuracy of cycle-counts – essentially moving from periodic to perpetual inventory.

Another function that many facilities perform is recording product data for tracking purposes. This is extremely important in pharmaceutical and food distribution facilities where information such as lot codes and expiration dates need to be recorded for manifesting as the product is picked. In many parts of the world, this information is mandated and must be reported to the customer and government authorities to enable a fast recall if a product is found to be tainted. Voice can simply ask the picker to speak the lot number or expiration date as a step in the picking process, or it can even direct a worker to pick a product with a particular lot code. This process can also combine voice with scanning.

In a similar manner, as with the experiences of Comercial Mexicana, voice can capture catch weights. This is used in grocery distribution to record the weight of a product as it is selected. Meats, for example, are packaged as cuts that do not all have a uniform size or weight. Since customers are charged by weight, capturing that information as the item is picked is essential to accurate manifesting and billing. The voice system either asks the worker to weigh the item and then to read the weight from a scale or, in the case of pre-weighed items, simply asks the worker to read the weight off of the package label as it is being picked. This information is then transferred to the warehouse management system, the transportation management system, and other enterprise and administrative software.

Value-added functions can also be controlled by voice. Clothing can be placed onto hangers, light manufacturing can be performed – just about any conceivable activity can be directed by voice and made more efficient in the process. These directions can be particularly important for DCs performing diverse value-added services for different downstream customers. For example, a clothing wholesaler may have some customers who want items on hangers and some who want the exact same SKUs folded. Sysco Guest Supply also uses voice for item consolidation. If the same item is found in multiple locations, based on the cube of the item and

the cube of the location, the system drives the user to combine the product into one location. Says Jim Makowski, "This creates various efficiencies for us....it not only creates open locations for spikes in inventory, but also affords us the opportunity to reduce the number of location selections when performing replenishments and other processes."

A global energy management company uses voice to improve accuracy for its kitting process (putting multiple parts into a kit) in one of its warehouse assembly lines. Kitting personnel operate in assembly-line function to kit parts for breaker boxes before final assembly. Working on a static conveyor, operators identify a unique order number, and the voice system directs the operators on which parts to pick and prompts for a location check digit to ensure accuracy. Since many of its breakers have a similar appearance, historically there were a number of picking errors. With the voice system, now once picking is complete within an operator's picking zone, he passes the assignment and tote to the next kitter in line and the process repeats itself with another order.

In the pilot implementation alone, errors were reduced by more than 50 percent and results noted in less than one month. Based on these results, the company implemented voice for additional product lines.

Another area ripe for voice is the packing station, where orders get packaged and moved out the door. This is particularly helpful for retailers and mail order forms requiring item customization, as well as those companies with a heavy internet business. Says Fox Racing's Robby Dhesi, "We will soon launch a packing station application for voice in our DCs. We think it is going to help us eliminate as much as 50 percent of unnecessary travel time, which will help us better meet our seasonal order demands with our existing workforce. Voice enables the highest level of accuracy and detail required for e-commerce, catalog and mail order operations to maintain top-level customer service. The superior accuracy also reduces the cost of returns." Dhesi also plans to eliminate scanners at the stations, thus reducing the amount of required hardware.

WHICH WAY DO YOU GO?

So far, we have examined many of the ways voice is being used today beyond picking in the distribution arena. Which begs the question – which is better, starting with voice across multiple workflows, such as Sysco Guest Supply and others have done, or transitioning from picking to other workflows, as many other companies have done? The short answer: *it depends.*

Either way, says Krupka of Burris Logistics, "You have to be ready for a certain amount of process reengineering. Voice will cause you to rethink and redesign your workflows, and that is a good thing – because voice provides a real-time window on your operation like no other distribution technology currently available. We aren't yet at the point where we are able to interleave tasks, because we are still in the middle of that reengineering effort. But once our processes are revamped, we'll be expanding voice to receiving and other workflows and we will be able to interleave between various tasks."

Because voice-enabled work started with the picking process, many companies that have realized the benefits of voice in this workflow are ready to move "beyond picking." A natural next step is to implement voice in the "adjacent" workflows of replenishment, stocking and putaway. When voice is used both in replenishment and picking, these two workflows can remain synchronized, replenishing pick slots during the pick and reducing shorts. When voice is also used in putaway, replenishment and putaway can be interleaved, saving labor as well as capital costs.

Using voice in cycle-counting is also a natural extension of the use of voice in picking. When cycle-counting is performed by dedicated personnel, it is typically an activity that involves the hands and eyes to a large degree, and therefore lends itself to voice. When cycle-counting is performed by pickers as part of a perpetual inventory process, using the same voice-based system that is used in picking is a logical decision.

Once voice has been established in picking, putaway, replenishment, and cycle-counting, outbound/loading is a logical next step. Since errors can be introduced in the loading process, using voice in loading can generate gains in both productivity and accuracy. Using voice in loading can also synchronize loading personnel with the pickers to ensure balanced workflows.

However, voice also is increasingly becoming the starting point for warehouse process redesign. When a DC makes the initial change from a paper-based system or implements a new WMS, many companies are choosing to voice-enable their entire facility with the initial project. In addition, since all the workflow processes are being designed or redesigned, developing these processes around voice provides a platform for cross-workflow productivity and accuracy gains both now and in the future. It also creates a balanced DC where workflows are synchronized with one another. In addition, workers need to learn only one new system or technology.

VOICE-ENABLING YOUR WAREHOUSE

Voice technology in the DC provides the transformational change needed to address today's key business challenges of controlling operating costs while improving customer service and enabling growth. As we have seen, voice can be used for a wide number of applications within a DC. But the real value is seen when voice is used for all or many of these functions. It creates huge gains in accuracy, assures high levels of productivity, reduces overall training time, and promotes safety, all at an attractive return on investment that is typically less than a year. And that cost comes down with each added application, since the same equipment can be used for all of the tasks if they happen on different shifts.

Once workers learn to work with voice, they can easily move to any functional area, where they simply follow the voice directions for that application. This common platform provides tremendous flexibility in planning, allows for fast customer response in filling orders including rush and emergency orders, enables savings through equipment re-use, and allows interleaving of tasks and switching workers between functions for higher productivity. The voice-enabled warehouse is the truly optimized warehouse.

In our final chapter of *The Talking Warehouse*, we will see how to make voice succeed for you. We will look at the characteristics of warehouses and DCs that would benefit the most from voice. We will explore what is required to deploy voice within a facility, and how to best integrate the technology with existing software and material-handling systems. We will also hear some lessons learned from those who have successfully implemented voice and are now reaping huge dividends since moving to voice-enabled operations.