Voice in the Warehouse: All Talk, or True Solution?

Building the Business Case for Voice Technology

Tom Singer, Tompkins Associates

July 2008

Introduction

Voice technology is so appealing today. It offers the promise of hands-free, eyes-free wireless access to the information needed to drive key warehouse processes. For warehousing applications, Voice can provide a fluid, natural user interface.

Before Voice, no other technology had a greater impact on the evolution of warehouse management systems (WMSs) than the wireless local area network (LAN), and mobile or Radio Frequency (RF) terminal. While they are popular with many organizations, RF terminals and barcode scanning do have some drawbacks. They require operators to use their hands to scan and key data. They also require operators to read instructions on terminal displays. For many operations, these activities disrupt the normal flow of warehousing and limit the benefits provided by the technology.

Despite these drawbacks, LANs and RF have provided the opportunity for many distribution operations to substantially increase accuracy, productivity, visibility, and control of core warehousing functions. In many ways, the integration of RF communications and barcode scanning into WMS solutions in the 1990s put information and data collection into the hands of warehouse floor personnel.

Like traditional RF-based barcode scanning terminals, Voice solutions center on a small, wireless computer that warehouse associates typically wear on a belt. The difference is that Voice delivers application instructions verbally through a headset and captures worker responses through a microphone – with no stopping to look at a screen, key in a quantity, or scan a bar code.

Voice Becomes Mainstream

Voice is not new in the warehouse. But until recently, it has been more of a niche player than a mainstream solution. This has started to change as the technology has matured and evolved. Voice is now poised to play a major role within the warehouse.

The key drivers in this movement are:

- **Proliferation of Wireless LANs**

  Warehousing has been at the forefront of the development of wireless LANs since the late 1980s. But early deployments of this technology were costly, custom propositions. The evolution of 802.11 standards propelled wireless LANs from custom to commonplace. Today, most top and mid-tier distribution operations use RF terminals with bar code scanners over 2.4 Ghz wireless LANs. Voice logistics solutions can share this common backbone with traditional RF scanning applications. Even operations that currently do not use RF scanning terminals are not put off by the prospect of installing an 802.11x (a,b,g) wireless network, since the technology has become commonplace in our increasingly mobile society.

- **More Powerful Mobile Devices and Standardization**

  The processing power of mobile computers and terminals has grown dramatically since the late 1990s. This, coupled with the continued evolution of Microsoft's family of mobile operating systems, has allowed Voice vendors to offer more powerful and robust solution sets. The Windows-Embedded product line provides hardware and software suppliers with a standard operating system to build upon. The old model of bundled proprietary hardware and software solutions has started to give way to applications that can run voice-enabled mobile computers offered by a multitude of hardware vendors. Voice application providers have also begun to embrace Service Oriented Architecture (SOA) and employ XML-based client software. Open standards generally mean more choice and better prices as well as protection for technology investments.

- **More Vendors and Solutions**

  The number of vendors offering Voice logistics solutions has steadily grown over the past few years. Most warehouse deployments come from vendors offering standalone application and Voice client software. The application software is interfaced to host WMS or legacy fulfillment systems through batch-like file transfers. In recent years, Voice logistics providers have begun to team up with WMS vendors offering a standard interface between the Voice vendor's client software and the WMS. Under this paradigm, a WMS package can support Voice functionality out-of-the-box, just like RF, with their Voice client communicating directly with the WMS application.

  Given the increasing interest in Voice and maturing of the technology, other types of system providers have started to offer voice-enabled solutions. Data collection systems vendors are supporting Voice in the systems which had been traditionally used as tool sets to provide RF functionality to legacy systems lacking mobile capabilities.

Automated order processing solution providers such as pick-to-light vendors have also started to voice-enable their application software. These new players are typically bundling established Voice solutions into their product suites.

- **Multi-modal Devices**

The traditional model for deploying voice within the warehouse centers on the use of voice-only terminals. These devices are basically wearable, wireless PCs without keyboard or display. In their initial incarnation, they were proprietary devices coupled to a single Voice vendor's software client and application. They can be accessorized by adding a bar code scanner. But their primary user interface for voice-only terminals is the headset and microphone. As the name implies, voice-only terminals can only be used for voice-specific applications. They cannot be used for functions that are not voice-enabled.

The adoption of Windows Mobile and Windows CE operating systems by both mobile computing vendors and Voice logistics solution providers has resulted in a new generation of multi-modal devices for warehousing systems. Mobile computing vendors have voice-enabled specific keyboard/scanner models through relatively inexpensive hardware and firmware variations. In addition, Voice logistics providers certify their client software to run on these modules. The result is a mobile computer that can be used either as a traditional RF scanning device or a Voice terminal.

By not being dedicated to a single mode, these devices provide prospective Voice users with flexibility that the old voice-only proprietary model could not offer. Multi-modal terminals can be shifted around the warehouse to meet changing work load demands, without worrying whether a specific function is voice or scanner/display based. They provide a vehicle that allows the user to choose the right input and instructional media for the job at hand. For example, the first half of a putaway task could be initiated by scanning a pallet label bar code. Instructions would then be provided verbally to the fork lift driver, who confirms the putaway by speaking the target location's check digits.

Voice-only terminals will continue to be popular in the near term. Many Voice applications still need more development to take full advantage of the multi- modal model. Some Voice solution providers price their multi-modal Voice clients so as not to erode their proprietary Voice-only terminals. But the trend is clear. The mobile computing industry is embracing a model where Voice is an inherent component of their solution set, and Voice solution providers will continue to build on this trend.

Just Promises, or Proven Potential?

More choice, greater flexibility, better price points, and a standard infrastructure – it is no

wonder that interest in this technology continues to build. But this increased overall awareness in the logistics world should matter little to any individual operation unless it can build a solid business case for the technology.

Voice offers the promise of improved performance and rapid return on investment. Trade journals, vendor product literature, and web sites are full of case studies and client testimonials that make a compelling case for the technology. Anyone who has worked with the technology knows that Voice has the potential to deliver on its promises.

But this doesn't mean that Voice is the right solution for every operation. While the benefits reported by one organization may be enticing, they may be much more limited or unobtainable for another operation. Also, implementation costs can vary significantly, depending upon legacy fulfillment or warehousing system, existing network infrastructure, and operational requirements.

So how does Voice actually stack up in the warehouse? Are vendor claims about benefits and quick ROI really valid?

Obviously, the answers to these questions will vary based on the nature of the operation considering the technology. Like any prospective technology investment, Voice implementation needs to be built on a solid business case in order to truly succeed. The depth and components of this justification can vary across operations. But it should always start with a thorough understanding of the operational and business requirements. It also needs to be based on a basic understanding of Voice's

- typical usages and alternatives.

- prospective benefits measured against the alternatives.

- key technology components and application integration.

- cost factors and implementation approaches.

Building a solid business case may require a fair amount of effort beyond this basic evaluation framework. Detailing benefits and costs to a detail necessary to adequately determine ROI takes work. But even if initial findings indicate that Voice is currently not practical for a specific operation, it will not be a wasted effort. While Voice may not be justifiable today, it may well be viable tomorrow for certain organizations.

Key Takeaways:

- The key drivers of Voice's increasing role in the warehouse are the proliferation of wireless LANs, the standardization of powerful mobile devices, the availability of more vendors and solutions, and the use of multi- modal devices.

- Voice can allow for more choice, flexibility, better price points, standardization of infrastructure, improved performance, and rapid return on investment.

- How Voice stacks up in the warehouse and the validity of vendor claims depends on the typical usages and alternatives, benefits measured against alternatives, key technology components and application integration, and cost factors and available implementation approaches.

Typical Uses and Alternatives

Voice is typically employed to support order picking within the warehouse. Grocery and Food and Beverage were the first industries to embrace Voice technology for picking.

Anyone who has worked in a refrigerated or frozen food pick module where gloves are standard equipment can appreciate the advantages of Voice's hands-free access. According to the Supply Chain Consortium's (www.supplychainconsortium.com) Benchmarking and Best Practices Review, over 85% of the respondents who stated that they were using Voice to support picking operations were in this industry segment.

But Voice has made significant inroads in other industry segments, including Automotive and Service Parts, Personal Care, Office Supplies, Healthcare and Pharmaceuticals, and Retail. While it is primarily employed for picking, some operations use it to support receiving, putaway, replenishment, and cycle count activities. Voice solution providers have recently begun to address functionality beyond picking in their standard product set. Picking remains a core focal point of interest for most Voice applications, given its proportion of overall labor activity in the DC and its direct impact on customer service levels. Typically, picking is the key component of establishing a business case for Voice in the warehouse.

The most commonly cited Voice alternatives are paper/label processing, RF terminals with barcode scanning, and pick-to-light.

- **Paper/label processing** is typically coupled with after-the-fact data entry using desktop terminals. Associates perform warehouse tasks off of pick lists, putaway labels, printed VAS instructions, and other paper documents. Upstream processes (such as how the information is sorted on the documents), and downstream processes, (such as scan and verify on a desktop terminal), directly impact paper/label

processing's performance and functionality.

Paper/label processing is a good fit for many warehouses, especially smaller operations with relatively straightforward transaction requirements. Even operations that rely on RF scanning for the bulk of transactions usually employ paper/label processing for some functions. It can be purely a manual proposition or part of an automatic flow, such as a label case pick-to-belt, where the pick is confirmed by an in-line conveyor scan.

- **RF scanning terminals** have been considered a prerequisite for larger, more complex operations. But RF scanning can be found in all different types and sizes of operations primarily due to direct support by most warehouse management systems. Even operations running non-RF enabled legacy fulfillment systems can turn to automated data collection software for this functionality.

  RF scanning offers some distinct advantages over paper/label processing. It can provide positive verification that the warehouse associate is at the right location or picked the correct SKU through a bar code scan or key entry. Work can be pushed out to associates based on location and task priority instead of handed out from a manually managed queue. Transaction data is captured in real time as associates perform tasks. Furthermore, RF scanning makes some functions like multi-order cart picking possible or more practical than paper/label processing.

- **Pick-to-Light (PTL)** remains a popular picking technology due to its ability to support high pick rates and ease-of-use. It is typically used in a zone-based, pick and pass flow where an associate scans a tote or carton bar code label. (The PTL software activates light displays for every location that shows the required quantity needed for the tote or carton.) The associate walks the zone, picking SKUs, and confirming picks by pressing display buttons. Pick quantities can be shorted or increased by button presses. Displays can also be provided to display SKU, order, or other relevant information. Some vendors even have LCD displays that show SKU pictures.

  PTL technology has a number of different variations. Instead of a full quantity display with confirmation buttons per location, a simple light indicator can be provided for each pick face with quantity shown and confirmed on a bay display. This configuration is generally employed in pick modules with slower moving SKUs. The technology can be used to support put-to-light packing, where an associate scans a case bar code and the PTL software identifies all the staged cartons requiring the case's SKU and associated quantity. Some vendors offer PTL "smart" carts where totes or cartons are associated with light-enabled cart slots. Associates push these carts through the pick module based on the location shown on the cart's light display. Once the location is confirmed through a wireless bar code scan, the PTL software illuminates the quantity needed for each slot requiring the SKU.

There are many other data collection and material handling technologies that are used to drive warehouse processes. But paper/label, RF, Voice, and PTL remain the most popular picking technologies. This observation is borne out by a recent Supply Chain Consortium survey *(Figure 1)*. So it should be no surprise that Voice vendors typically highlight their wares against the other three pick methods.

Figure 1. Picking Technologies Used by Survey Respondents, as Reported by the Supply Chain Consortium

Key Takeaways:

- Voice is primarily used in picking operations, which generally makes picking the key component for a business case for Voice. However, some use it for receiving, putaway, replenishment, and cycle counting.

- The most popular alternatives to Voice are paper label processing, RF terminals with barcode scanning, and pick-to-light.

Weighing the Potential Benefits of Voice

Vendors point to a variety of potential benefits for employing Voice within the distribution center. They typically provide metrics for these improvements based on case study data from their client base. Moreover, the benefits of Voice and associated metrics have been well documented in numerous trade journal articles and white papers. These reported improvements or reductions are usually impressive, but must be viewed within the context of *before* and *after* points. They must also be examined against the nature of the operation, product being handled, and systems involved. However, testimonials do provide a general indication of what Voice can do in the warehouse.

Typical Vendor Data

While classifications and measurements may vary between case study and web site, they fall along the following lines:

- Increased productivity and pick rates

- Reduced errors and increased accuracy

- Improved throughput and fill rate

- Reduced supply costs

- Improved control and visibility

- Decreased training time

- Improved safety

- Reduced damage and breakage

- Faster worker training

- Enhanced worker satisfaction

Voice vendor web sites provide case study data quantifying many of these benefits, especially productivity and accuracy gains. Reported productivity increases usually range from 8-40%. Occasionally, higher increases may occur. The average case study cites gains between 10-20%. Vendor marketing collateral and industry reports typically mention a slightly higher than expected improvement, ranging between 15-25%.

Most vendors report one or more customers who have at least doubled pick rates. Generally, the featured operation for which the performance measurement is provided is moving from paper or label-based picking to Voice, with a few case studies based on implementations replacing RF scanning or pick-to-light.

Accuracy rates typically cited in these case studies are at least 99.5%, with most reporting higher rates. Corresponding reported reduction in pick error rates range from 80-100%. Some studies detail significant cost savings in supplies (moving from label to Voice picking) and increased fill rates due to reductions in mis-picks.

Figures for other benefits, such as improved safety and reduction in breakage, are not generally reported. Given Voice's hands-free and heads-up processing flow, these benefits make intuitive sense. Vendors generally showcase customers who obtained a return on investment (ROI) between 9 and 12 months.

Voice appears to be an attractive investment proposition in the warehouse. But are the numbers cited on vendor web sites realistic for a specific operation? While there is no

reason to doubt the veracity of vendors' numbers, they must be viewed in the context of the starting point and processes involved.

Measuring Gains in Productivity and Accuracy

Potential productivity gains can be quite significant for an operation moving from paper to Voice. In general, these gains are due to a number of factors beyond the hands-free flow of Voice, including:

- Changes in pick process, such as moving from discrete order picking using paper pick lists to multi-order cart picking, using functionality provided by the Voice application software.

- Reduction in personnel needed for post-pick checking, packing, and auditing, due to positive pick verification of Voice over paper picks.

- Real-time information on inventory levels, order status and picker transaction rates provided by the Voice application software.

Depending on the functionality provided by the underlying software, the above factors generally do not play a significant role when comparing RF Scanning to Voice. From a productivity perspective, the comparison between the two pick methods centers more on the hands-free nature of Voice.

The Tompkins Associates White Paper, *Order Picking for the 21st Century, Voice vs. Scanning Technology*, documents a 2003 implementation of the Vocollect Voice solution at Associated Wholesale Grocers (AWG). The implementation covered five pick areas: dairy, dry, freezer, meat, and perishables.

Figure 2 shows the productivity gains after Voice was installed. The two areas, Dry and Freezer, which were previously supported by a paper-based process, experienced modest gains. The areas previously supported by RF scanning saw much higher increases. This is not surprising, because RF scanning can be more disruptive to picking flow, since it typically requires the user to scan or enter information at multiple points. Equally understandable are the higher gains in the refrigerated Meat and Dairy areas, where RF scanning terminals can be more difficult to handle.

Area	Old Pick Method	Picking Productivity Gain
Dry	Paper	3%
Freezer	Paper	4%
Produce	RF Scanning	8%
Meat	RF Scanning	12%

Dairy   RF Scanning   15%

Figure 2. Productivity Increases after Voice at AWG

While the relatively modest gain in picking productivity for Voice over paper may be expected, other factors should be considered when comparing paper to other picking technologies. Paper generally requires post-pick data entry, either at a packing or clerical key entry. Overall productivity gains of moving off of paper need to account for reduction in these efforts. Since picks are not systematically verified as each line is processed, errors are more likely to occur; and correcting these errors requires additional labor. Paper also requires manual management that covers preparation, assignment, and post-pick processing. This all entails additional direct and indirect labor that should be considered when quantifying prospective labor productivity gains.

Voice versus RF Scanning comparisons should also account for the different types of the RF Scanning devices, generally categorized as handheld, wearable, or truck mounted. Handheld terminals typically require users to holster or set the device down during certain steps in a process. This can add to the overall time to complete a transaction. Wearable units are worn on arms or attached to belts. These lightweight devices capture bar code data through "ring" scanners worn on the index finger. Truck mount units are mounted on material handling equipment such as reach- and order-picker trucks and motorized pallet jacks. Truck mounted devices generally capture bar code data through tethered scanners. While wearable and truck mounted units do not require users to pick up or lay down the device, warehouse associates still must read the display and key data at certain steps in a process, potentially slowing down overall transaction time.

Pick-to-light vendor web sites also claim similar benefits for their light-based solutions. Like Voice, productivity and accuracy are typically the cornerstones of any pick-to-light business case. Some vendor web sites cite 4- or 5-fold productivity improvements over paper-based picking, with individual pick rates approaching 450 lines per hour. While these numbers may seem high, pick-to-light is generally acknowledged as providing the highest pick rate potential of the four picking technologies when pick densities are relatively high.

Some sources assert that Voice provides a greater accuracy potential than RF Scanning and Pick-to-Light. All three technologies can provide significantly lower pick errors than paper, since they all require positive real-time confirmation of the pick. However, some studies report lower error rates with Voice than the other two methods, due to the freeing of hands and eyes from data entry steps.

RF Scanning does require the picker to break the flow of the process to perform scans, read displays, and key quantities. Arguably, these breaks in flow can interject errors into the process. But pick-to-light only requires the push of a button to verify the pick. The actual speed of pick-to-light may generate slightly higher error rates than Voice in certain situations, as pickers may concentrate too much on speed at the expense of paying attention

to the pick task at hand.

Quantifying Benefits

Case studies can provide a good general indication of the potential of Voice. But they tell stories for specific operations, making them less applicable to any individual distribution center. The potential fit of Voice or any other picking technology is dependent on a variety of underlying factors, including:

- Order profile – lines per order, and units per lines.

- SKU weight and size.

- Pick container weight and size.

- Travel distance between picks.

- Pick line layout and product accessibility.

- Special data capture requirements such as lot, batch, serial number, or catch weight.

- Workforce composition, including percentage of temporary workers.

- Growth potential and need for flexibility.

- Functionality of the supporting software application.

Since these factors can vary across operations, building a business case for Voice on the benefits obtained at other sites can be risky. Benefits can certainly be quantified by conducting pilot tests. On the other hand, pilot programs are generally costly and impractical. This leaves two viable alternatives when quantifying benefits: 1) Using case study data and assumptions, or 2) Developing engineer-based analysis of anticipated gains.

- **Case study data**

  Certainly the risk involved in using case study data and generalized assumptions to quantify benefits is a function of how much the target operation differs from the case

study operations or falls outside of the "norm" that is the basis for the general assumption. For many operations contemplating Voice, this should be a perfectly acceptable risk, especially if conservative numbers are used. Using 10- 12% as the anticipated labor productivity increase in moving from paper or RF scanning to Voice is generally a good rule of thumb. But it does not account for variances in operational flow, layout, product, personnel, and legacy systems.

- **Engineer-based analysis**

  Quantifying potential benefits through an engineer-based analysis can account for these variances. This approach breaks down the elemental processes and steps for current and prospective processes. It can probably be best appreciated in the context of developing expected pick rates from predetermined elemental tasks and associated time. Employing this approach for quantifying potential pick rates allows for comparisons between technologies and process flows, as well as accounts for variations in the above factors, if properly done. It is a method requiring specific skill sets in order to produce reliable results and is generally performed by an industrial engineer.

  Figure 3 shows an example of the results of a predetermined time element analysis performed for a Tompkins client. It summarizes anticipated case pick rates in cases per hour between paper, RF scanning, and Voice in a refrigerated pick module. Detailed analysis for Voice picking appears in *Figure 4*. The analysis was developed using time sampling of existing paper-based pick processes, as well as elemental step evaluation for the potential use of RF Scanning and Voice. The results show pick rate increases of 6% and 12% respectively for moving from paper and RF Scanning to Voice.

Pick Technology	Cases/Hour
Paper	196
RF Scanning	184
Voice	209

Figure 3. Anticipated Case Pick Rates in Cooler Module

Once again, the paper pick rate increase does not account for post-picking data entry and verification. Also, the legacy system employed at this operation produced paper pick lists during nightly runs. Order changes that occurred after the pick lists were printed required additional processing.

Picking – Cases	Model	Index	Freq	Factor	Total TMU's	Total Hours	Cases / Hr
Walk to Vehicle	A	1	1	100	100	0.001	

Element	Code					
Start and Park	S	3	1	100	300	0.003
Transport	T	3	1	100	300	0.003
Load Empty Pallets	L	3	20.96	10	628.8	0.006
Transport-Directed to Location by Voice	T	3	20.96	100	6288	0.063
Load Case onto Pallet						
Action Distance	A	3	1048	10	31440	0.314
Body Motion	B	3	1048	10	31440	0.314
Gain Control	G	3	1048	10	31440	0.314
Action Distance	A	3	1048	10	31440	0.314
Body Motion	B	3	1048	10	31440	0.314
Placement	P	6	1048	10	62880	0.629
Action Distance	A	3	1048	10	31440	0.314
Place Label on Case						
Action Distance	A	1	1048	10	10480	0.105
Body Motion	B	3	1048	10	31440	0.314
Gain Control	G	3	1048	10	31440	0.314
Action Distance	A	3	1048	10	31440	0.314
Body Motion	B	3	1048	10	31440	0.314
Placement	P	3	1048	10	31440	0.314
Action Distance	A	3	1048	10	31440	0.314
Load – Pick up Pallet	L	10	20.96	100	20960	0.210
Transport – Travel to Next Location	T	1	148	100	14800	0.148
Transport back to Dock w/ Pallets	T	3	20.96	100	6288	0.063
Stop Vehicle	S	6	1	100	600	0.063
				Total:	500,905 5.009	209.2

Figure 4. Sample Voice Case Picking Elemental Analysis

Figure 5 shows an example in which pick-to-light and Voice were analyzed for different pick modules and order types. The software solution being evaluated for this operation provided pick-to-light, RF scanning, and Voice picking functionality. The summary results in Figure 5 show that pick-to-light provides a significantly high pick rate for Store Orders, especially in the Carton Flow module. But Voice and pick-to-light have compatible pick rates for Service Orders in Shelving. These rates were incorporated into a cost benefits analysis that recommended deployment of both technologies in separate pick modules.

Store Orders Service Orders

	Discrete Pick-to-Light	261	132
Carton Flow	Discrete PTL Both Directions	278	162
	Batch Voice Picking	186	147
Shelving	Discrete Pick-to-Light	178	120
	Batch Voice Picking	141	111

Figure 5. Anticipated Pick Rates in Lines/Hour

An engineering-based approach can be used to quantify other benefits. However, assumptions may have to be made in certain situations. In this case, sensitivity analysis can gauge the impact of varying these assumptions on the results. Knowing how the software application truly functions is critical to quantifying realistic benefits. It may also factor in the cost portion of the analysis in the event that software modifications are required. Appreciation of how Voice application software performs for any specific operations starts with a general understanding of its key technology components and integration to warehouse systems.

Key Takeaways:

- Vendors use case study data to report improvements, but these reports must be considered in the context of *before* and *after* points, as well as the nature of the operation, product being handled, and the systems involved.

- Productivity and accuracy are the cornerstones of a business case for Voice.

- Using a pilot program to quantify the benefits of Voice may be too costly; instead, consider using case study data or engineer-based analysis as viable alternatives.

Technology Components & Application Integration

For the most part, Voice logistics solutions share a common hardware and software architecture. At first glance, they appear to look and work similarly. But under the covers, there can be some significant differences between vendor solutions. Furthermore, practical choices in prospective vendor solutions for any specific operation may be limited by a number of factors, including the WMS package used.

For many operations, Voice can be approached as a shrink-wrapped application in which vendor quoted costs and performance will typically match the results experienced.

Occasionally, the underlying architecture and integration to warehouse applications of a specific solution significantly impact cost and performance.

Making assumptions about how Voice works (or any other data collection technology) in any particular situation – based either on generalizations or how the solution performs at other operations – creates the risk of unexpected and potentially unpleasant results during implementation. Organizations can minimize this risk by taking the time and effort to understand Voice's basic components and how it interacts with other warehouse systems.

In many ways, Voice employs a similar technology infrastructure as RF Scanning. It is a distributed technology that uses an $802.11x$ standard wireless LAN that supports communications between client mobile computers and backend servers. The mobile devices typically employ the Windows CE operating system. Client software running on these devices manages data presentation and input services, which in the case of Voice means speech recognition and text-to-speech functionality. Servers provide business application and database functionality, as well as Voice client administration.

Client Components

Within this distributed framework, Voice solutions can vary significantly in how these components function both from the client and server perspective. Client speech recognition components are either speaker-dependent or speaker-independent. Speaker- dependent requires users to train the system to recognize the specific nuances of their voices. This process involves the system prompting the user to repeat digits and terms. Individual voice templates are stored on a management server and downloaded to the mobile devices as needed. Generally, it takes 15-20 minutes for a user to create his/her voice template.

The speaker-independent approach does not require the user to train the system. This is the method employed by voice-enabled telephone customer service applications, where callers respond verbally to system prompts. Most Voice vendors support only one approach. Those offering a speaker-dependent solution generally claim that the speaker- independent approach is not as dependable in recognizing responses in the relatively noisy environment of most warehouses. Also, the speaker-independent method can be challenged by regional dialect variances. Speaker-independent solution providers generally contest these claims. Comparing reference site environments to the targeted operation can help sort through these conflicting statements. Other factors such as headset quality can impact Voice performance and reliability in the warehouse.

Voice recognition works best in the warehouse when user responses are limited to short distinct phrases and digits. Location verification is generally done by repeating two digit numeric check digits associated with each location. Lengthy responses can present challenges, both from recognition and performance perspectives. Voice may be an excellent tool for capturing check weights, but bar code scanning may be a better choice for recording serial numbers.

Voice solutions can also vary in how their client software interacts with backend application servers. Most employ operation specific programs or task files that are downloaded to the voice-enabled devices. Communication between the client and application servers is controlled by these code sets, as well as support for client-side functional processing.

Some Voice solutions use a model that dispenses with client-side voice specific application code. These solutions treat voice as another input/output stream no different than text displayed or entered on a handheld computer. Input and output mapping is handled by server-based processes. Client-based software handles the Local presentation and data capture functions. Performance factors and application integration typically govern which approach is employed.

Server Components

Voice clients communicate with backend servers for application processing and database services. While some functionality may reside on client devices, most data validation and processing logic occur on application servers. Higher level systems, such as a WMS or order management system, provide order and inventory data for execution.

There are two basic approaches for integrating Voice into a WMS or order fulfillment system: *direct interface* and *standalone application*.

Under direct interface, client software exchanges information in real time directly with the WMS or higher level host system. This is done through a predefined set of application programming interfaces or service messages that each side can use to send or receive data from the other side. A number of Voice vendors facilitate this approach by publishing a standard library of message transactions. While most top-tier WMS solutions support a direct Voice interface, many WMS packages do not. Moreover, WMS vendors that support a direct interface typically only do so for a single Voice vendor.

Most Voice logistics vendors provide standalone application software capable of supporting core warehousing operations much like a lower-end WMS package. Order and inventory data is downloaded from the higher level host system. All transaction processing occurs on the standalone application, with resulting pick confirmation and inventory data uploaded to the host system. This approach provides the potential for WMS integration through a relatively limited set of interface points. For example, pick demand can be downloaded upon waving, and pick responses uploaded after the transaction has been completed.

Building this batch-like interface may be more economical in certain situations than constructing a real time direct interface. Furthermore, it allows operations using order or inventory management systems to take advantage of warehousing functionality that may be unavailable in their legacy systems. For example, implementing a Voice logistics standalone application may allow an operation to move away from discrete order picking to a more efficient zone, batch, or multiple order picking process.

A direct interface is more attractive from a cost and performance basis if it is already

supported by the WMS vendor. These interfaces provide "out-of-the-box" Voice functionality that requires no additional programming or development – provided the functionality meets the specific requirements of an operation. Generally, this is the case.

But in certain situations, both Voice client and WMS must be modified to meet requirements. Care should be taken when comparing specific requirements to the base Voice functionality supported by a WMS vendor. It should never be taken for granted that WMS Voice and RF Scanning functionality work exactly the same.

Key Takeaways:

- After selecting a Voice vendor solution, cost and performance can be impacted by the underlying architecture and integration to warehouse applications – a risk that can be mitigated by understanding Voice's basic components and integration with other systems.

- Client speech recognition components are either *speaker-dependent* (requiring users to train the system to recognize their speech) *or speaker-independent* (no training is necessary).

- There are two approaches for integrating Voice: *direct interface* and *standalone application*.

Cost Factors and Implementation Approaches

Developing a cost estimate for implementing Voice in a distribution center is a somewhat similar exercise to determining the cost to install RF Scanning or any other Auto-ID technology. Costs may be broken down a number of ways, generally categorized as hardware, software, support, professional services, and customization. *Figure 6* shows a sample cost estimate developed for a single site implementation. Specific line items can vary, depending on the solution to be implemented and the existing infrastructure..

	Sites	FTEs	Shifts	
	1	70	3	
Item	Qty	Unit Price	Extended Price	Total
Voice-only Terminals – Wi-Fi Radio (+20% spares)	28	4,000	112,000	
Battery (1 per terminal)	28	200	5,600	
Charger – 5 Terminal	6	900	5,400	

Charger – 5-battery	6	900	5,400	
Wall mount charger	6	400	2,400	
Terminal cover (1 per terminal)	28	75	2,100	
Headsets and Belts – 1 per person +20%	84	300	25,200	
PC Connection cable – 1 + spare	2	100	200	
Training Device & Listening Kits	1	900	900	
Wireless listening & Maintenance kit	1	950	950	
Windscreen headsets (bag of 25)	1	50	50	
Hardware Total				$ 160,200
Software per terminal	28	600	16,800	
Application software – server based	1	10,000	10,000	
Software Total				$ 26,800
1 year Express depot	1	5,000	5,000	
1 year Support plan	1	1,000	1,000	
Support Total				$ 6,000
Pre-Implementation Visit	1	2,000	2,000	
Pre-Implementation Visit – T & E	1	3,000	3,000	
Professional Services (mgmt * Imp)	10	2,000	20,000	
Professional Services – T & E	1	3,000	3,000	
Vendor Professional Services Total				$ 28,000
Custom Development	20	1,500	30,000	
Custom Development Total				$ 30,000
Total Estimate				$ 251,000

Figure 6. Sample Voice Cost Estimate

Hardware includes Voice terminals as well as accessories such as headsets, battery chargers, and training devices (if a solution is based on speaker-dependent voice recognition). Allowances should be made for spare equipment. If the target facility does not have the necessary $802.11x$ wireless LAN coverage, then equipment and installation line items need to be specified for the RF network, including cabling. A server will be required for Voice management software, and additional servers may be needed for application and database software.

Software includes Voice vendor client and server license fees. It can also entail server operating system and database licenses. Extended warranty and support contracts should be detailed for hardware and software. Generally, vendors provide several annual support plans tailored to meet a variety of customer contingency preferences. Advance exchange options, where replacement units are immediately sent via expedited delivery upon report of

an original unit failure, represent the high end of available equipment support.

Professional services are typically required to support a Voice implementation. These fees cover vendor gap analysis to assess if any modifications are needed, installation and configuration support, training, and go-live support. If the solution features a direct WMS interface, additional professional services may be required from the WMS vendor.

Customization may also be needed if a solution features a direct WMS interface. Solution providers can vary by the package implemented and equipment used. Voice hardware and software may be available directly from the Voice vendor or from a vendor-authorized reseller. WMS vendors that provide a direct Voice interface may resell Voice enabled equipment, accessories, and software. They may also require their customers to purchase all required Voice components and services directly from the WMS vendor.

Some resellers provide extended integration services and are capable of performing modifications. Vendors that sell voice-only terminals usually bundle the cost of the Voice recognition client into the unit price. However, this cost is typically additional when purchasing a multi-modal terminal from vendors like LXE, Motorola, Intermec, and Psion Texlogix.

Developing the cost components for a business case or budget takes a certain skill set. Vendor quotes can be solicited to aid the process, but vendor provided-numbers should be critically reviewed and challenged. Furthermore, vendor numbers are only as good as the input provided by the prospective customer. If needs are not adequately presented, any estimate will probably be suspect.

As with any logistics systems project, there are always potential unknowns lurking. And these unknowns are typically addressed by adding a contingency factor to the cost equation. The biggest unknown is likely to be the need for custom development. Few vendors are going to suggest a high customization cost on an initial quote. In all fairness, it takes a full gap analysis and design process to determine how much customization, if any, is required. Vendors tend to be relatively optimistic up to a point.

Implementing Voice in the warehouse is an actual project and needs to be treated as such. It requires a project management framework and sufficient internal support resources to be successful. Operations must actively participate in design, testing, and training activities. IT resources must be provided to support installation, technical configuration, systems administration, and integration testing. If a standalone Voice application is employed, IT will probably have to devote resources to support the download and upload of files between the internal host system and Voice application.

Additional development hours may be required for enterprise and warehousing system customization to support the interface. Regardless of the support plan purchased, sufficient internal resources – both operations and IT – need to be devoted during implementation to take full ownership after go-live.

Key Takeaways:

- Categories of costs are hardware, software, support, professional services, and customization.

- No matter what support plan is purchased, internal resources from operations and IT must be devoted for design, training, installation, technical configuration, systems administration, and integration testing.

Conclusion: Moving Forward with Voice

Voice is not for every warehouse. There are plenty of facilities where paper-based processing is the optimal technology. RF scanning may provide better functionality and cost effectiveness in the long run. Other technologies may yield superior performance and return on investment.

However, the benefits cited in numerous Voice case studies are real and may be obtainable for any individual operation. Voice has moved beyond cutting edge to become an established warehouse technology. Any distribution operation concerned with improving productivity, accuracy, and throughput should give the technology serious consideration.

This should start with the realization that Voice is not a mutually exclusive proposition in the warehouse. Many operations that use Voice employ other technologies such as RF scanning and pick-to-light. What it boils down to is selecting the right tool for the job.

Managers of distribution operations need to approach any process or system improvement project from this perspective. Voice is merely one of the technology tools to be considered, and developing a sound business case that looks across available tools is a necessary first step.

Building a business case for Voice or any other technology in the warehouse requires careful delineation and quantification of benefits and costs. It entails an ability to detail current processes and requirements, map how these processes will change, and plan how requirements will be supported using the new technology. Some key factors to keep in mind when evaluating Voice for a particular warehouse operations are:

- **Keep the proper goal in mind** – The objective of any evaluation is not to figure out how to get Voice into the warehouse. It should be about selecting the best tool for the job.

- **Employ an evaluation approach appropriate to the situation** – The details and

depth needed for a successful evaluation are dependent on the current operation and systems. For example, an operation already using a WMS solution may want to consider using the package's direct Voice interface in a particular pick module that is currently supported by RF scanning. Costs, application, and integration components in this situation are much more concise than a paper-based operation without a WMS that is being compelled to significantly expand its capacity. The former may be able to get a relatively high level review, but the latter needs a comprehensive analysis.

- **Do your homework** – Operations managers do not need to become experts in the technology to consider its use. However, anyone evaluating Voice needs to know enough about its usage, alternatives, benefits, components, cost structure, and integration to make an informed decision. While Voice and WMS vendors can provide guidance and support in developing a business case, any organization contemplating the technology must be prepared to critically challenge its applicability within its distribution center.

- **Put together the right team** – Evaluating, implementing, and using Voice are multidisciplinary propositions. The success of any Voice evaluation project is contingent on putting together a cross-functional team that represents management, operations, and IT. Since it may entail a significant investment, finance may also be needed to help frame the business case approach. If adequate internal resources are not available or the evaluation is inherently complex, consider retaining the services of a third party consultant.

- **Be realistic and above board** – The ability to adequately state benefits and costs is the crux to any successful evaluation of a technology or system in the warehouse. However, assumptions and estimates are an inherent component of even the most structured evaluation process. No mater how scrupulous an organization is in its process, there is always the potential of some unknown factor compromising the end results. Some operations respond to this risk by being conservative on benefits and factoring in a contingency line item on costs. Others bracket minimum, expected, and optimistic savings/gains by benefit. Regardless of the approach employed, any operation evaluating the technology needs to occasionally step back and question whether the numbers being employed are realistic.

- **Treat your business case as living document** – Be prepared to live by the business case you develop. Track its performance during implementation and beyond go-live. Measure whether the anticipated ROI was achieved in projected timeframe. Many organizations do not perform post go-live assessments of their systems projects for a variety of reasons. This is wrong. Even if a project has missed its mark, knowing the root causes for the situation can present an opportunity to change course.

The expansion of interest in Voice and the proliferation of solution providers is not a fluke or hype. Voice has a real role to play within the warehouse and is rapidly becoming a mainstream technology. While it may not be viable in the near or even long term for many operations, many others stand to gain from its employment. The first step in this process is determining how it stacks up within your warehouse. Given the evolutionary aspect of Voice technology and applications, this is not a static proposition. If the technology is not a good fit today, it may be eminently viable tomorrow.

About Tompkins Associates

Tompkins Associates designs and integrates global end-to-end solutions for companies that embrace supply chain excellence. For more than 30 years, Tompkins has evolved with the marketplace to become the leading provider of global supply chain services, distribution operations consulting, technology implementation, material handling integration, and benchmarking and best practices. The company is headquartered in Raleigh, NC. For more information, visit www.tompkinsinc.com.

Innovative, practical solutions that improve your supply chain performance and produce value-based results.

References

The Supply Chain Consortium is the premier source for supply chain benchmarking and best practices knowledge. With more than 200 participating retail, manufacturing and wholesale/distribution companies, the Consortium sponsors a comprehensive repository of 17,000-plus benchmarks complemented by search capabilities, online analysis tools, topic forums and peer networking for supply chain executives and practitioners. The Consortium is led by the needs of its membership and an Advisory Board that includes executives from Campbell Soup, Hallmark Cards, Ingram Micro, Mervyn's, Molson Coors Brewing Co., Target, The Pep Boys, and Coca-Cola Co. To learn more about how your company can become a member of the Supply Chain Consortium, contact John Foley, 919-855-5461 or visit www.supplychainconsortium.com.

Contact Information:

To learn more:

Tom Singer, Principal
Tompkins Associates
tsinger@tompkinsinc.com

To learn more about best practices in order lead time and variability or the resources available through the Supply Chain Consortium:

Bruce Tompkins, Executive Director
Supply Chain Consortium
btompkins@supplychainconsortium.com

Chris Ferrell, Principal
Supply Chain Consortium
cferrell@supplychainconsortium.com

Tompkins Associates
6870 Perry Creek Road Raleigh, NC 27616
www.tompkinsinc.com
United States Canada Europe Asia