

## SPEECH SOLUTIONS AND v10 FEATURES OVERVIEW

### NEW FEATURES THAT ENABLE SPEECH

- ☑ **Screen Recognition and Reformatting.** A better way to integrate speech and automate existing terminal emulation (Telnet/SSH) screens.
  - **It starts with Screen Recognition**
    - **Define what's unique about a screen.** Screen Recognition allows for multiple criteria – including that something IS NOT on the screen. (i.e., "Pick" is in row 2, col 4 AND "Loc:" is in row 3, col 6 AND "ERROR" is NOT in row 16, col 1.)
    - **Define values on the screen.** Static text like titles and labels, or changing values like the location or quantity.
  - **Add actions to perform**

**Define instructions or actions.** Actions can be added to be performed the first time a screen appears or or upon returning from later screens (we can even tell that you returned from a screen using a cancel command vs. completing normally). Examples of automation commands:

    - **Say** something to the user using "Stay-Linked Speech" to control CTGTALK (optional add-on).
    - **Accept speech** recognition input using CTGTALK on the device.
    - Turn off the scanner
    - Send a message to the user
    - Force the device to beep – or beep a number of times.
    - Define automatic responses when a screen is recognized.
    - **Any combination of the above.**
  - **Reformat a recognized screen** (Optional)

**Redefine the size and content of a screen.** Create a newly resized screen to fit the device screen size. Define which values to show and where to place them on the new screen. Create new labels and commands.
- ☑ **New Stay-Linked SpeechDirect – Powered by the Nuance Engine**

Stay-Linked's newest offering is SpeechDirect™ - speech output for terminals. Not intended for speech recognition applications, it's a perfect fit when you want to add verbal instructions and prompts to existing applications without the complications of full speech recognition. (See the next page for more information.)
- ☑ **Better Integration with CTGTALK speech recognition engine**
  - **Speech engine interface improvements.** We've improved the interface with the speech engine to make it faster and even more reliable.
  - **Stay-Linked now resells and supports CTGTALK.** We're proud of our customer service and support. We now support the speech engine integration with Stay-Linked so you only have one company to call.
  - **Latest version of the speech engine.** Version 3 implements better recognition rules and speech gain control. The Grammar file that defines valid speech input is better and faster.

## A DIFFERENT APPROACH TO IMPLEMENTING SPEECH

Now there's a way to get **80% of the benefits** of speech for **less than 20% of the cost!**

### **Introducing Stay-Linked SpeechDirect™:**

- **Cheaper** up-front costs, **easier** to implement, and **faster** to production.
- **No changes to WMS/ERP** required. The Stay-Linked Server and Administrator Console do all the work.
- Proof-of-Concept (POC) in **HOURS instead of months**.
- Can be implemented by your own **internal staff** – not consultants.
- All configurations and commands are **centralized** in the Stay-Linked Server.
- Stay-Linked SpeechDirect runs on **"multi-modal"** devices — interchangeably use the scanner, keyboard, or screen taps for input.

### **"Speech Solutions" have 2 very different components:**

- **Text-to-Speech (TTS) – A quick and easy way to get started.**
  - Our suggestion is to **focus on TTS** – where **80% of the benefits** come from.
  - We call this **"Speech Direction"** – the WMS displays information on the device, and **audibly directs the user** where to go, what to pick, and how many.
  - **ROI** from Speech Direction is **quick** because **no user training is required**.
  - **80+% improvement in accuracy** when they see – and hear – the instructions. (Going to the right place and picking the right things the first time.)
  - **Headsets are cheaper** if you aren't paying for noise-cancelling microphones.
  - Your company already uses barcode scanning for input—continue to use and take advantage of your existing systems. Besides, scanning is faster than speech.
- **Speech Recognition-to-Text (STT)**
  - Implementing STT tends to extend the adoption cycle—and it may not be a substantial benefit to the company.
  - Training staff to properly use Speech Recognition takes time.
  - "Best Practices" must be used to improve speech recognition. Failure to adhere to the rules is the single-most common mistake, and causes the most recognition failures—which ultimately results in failed implementations.
  - Stay-Linked Speech enforces the use of Best Practices – which the others don't.

**Stay-Linked SpeechDirect** includes TTS only. By focusing on TTS – and not on STT – you won't be comparing other products that **MUST** use speech input. **Using only TTS provides you with a practical speech assisted solution.** And if you want Speech Recognition, the full Stay-Linked Speech™ is still the best solution for complete, accurate, multi-modal input that includes speech recognition.

**Only with Stay-Linked!**

## Terminal Emulation and Speech – Done Right!

### **SPEECHDIRECT: MAIN BENEFITS**

- **Inexpensive** – About \$600 list price per device. (includes Stay-Linked terminal emulation, terminal management, and SpeechDirect licenses and a headset)
- **Fast** – Proof-of-Concept in *hours* – not *months*!
- **Easy** – Full application speech output enabled by your own staff in a couple of days. (Assuming about 8 – 10 screens in the application)
- **Reliable** – All of the rules and configuration run, and stay, on the host system—which is your most reliable system.
- **No staff training** – Since it's output only, just give them the headset and they start hearing what's important.
- **It can't get much easier or quicker!**

### **FULL SPEECH RECOGNITION: MAIN BENEFITS**

- **Inexpensive** – About \$1500 list price per device. (includes Stay-Linked terminal emulation, terminal management, and CTGTALK licenses and a headset)
- **Fast** – Proof-of-Concept still in *days* – not *months*!
- **Easy** – Full application speech enabled by your own staff in a few days. (Assuming about 8 – 10 screens in the application)
- **Reliable** – All of the rules and configuration run, and stay, on the host system—which is your most reliable system.
- **It Works** – Certified and supported by Stay-Linked. And you know we won't put it out unless it works—and works well!
- **Training is Easy** – Our speech product doesn't require "profile training" – making it quicker for users to learn how to use speech recognition.

## **ADDITIONAL BENEFITS**

### **Stay-Linked Integration:**

- Host-Based Screen Recognition – Where all the power is!
- Creation and configuration of speech is integrated right into Stay-Linked's PC-based Administrator Console.
- Rules and commands are NOT on the devices – everything stays on the host and are easily managed using the Stay-Linked Administrator.

### **Speech Recognition – the ctgTALK Engine:**

- Speaker Independent Recognition – No up-front "Training" of the speech engine is needed.
- The speech engine "Adapts" to the speaker as they use it to adjust for changes in pronunciation and account for user fatigue.
- Subtraction Filtering (noise cancelling) *IN ADDITION* to microphone's built-in noise cancelling capabilities.
- Gain control automatically adjusts to speaker's voice level.
- Absolute Threshold adjusts for what is said above background noise.

## **COST & IMPLEMENTATION COMPARISON** **Full Speech Recognition Option**

	<b>Stay-Linked &amp; CTGTALK</b>	<b>Traditional Server Solutions</b>
<b>WIRELESS NETWORK</b>	Simple RF	Complex / high bandwidth
<b>DEVICES</b>	Muti-Modal scanners - \$1K - \$3K	Dedicated/Single-Function – \$3K+
<b>LICENSES PER DEVICE (List Price)</b>	\$250 – Stay-Linked Terminal Emulation plus \$995 (Speech) OR \$250 (SpeechDirect)	> \$1,500
<b>HEADSETS</b>	\$50 - \$100	\$150 - \$250
<b>APPLICATION</b>	Use current TE Apps	\$30K - \$50K For front-end app
<b>IMPLEMENTATION</b>	Internal staff - ~\$3K <b>(5 – 10 DAYS)</b>	\$30K - \$50K <b>PLUS</b> internal staff time <b>(2 – 6 MONTHS)</b>
<b>10 USERS – COSTS FOR INITIAL IMPLEMENTATION</b>	Full Speech: ~ \$15 K SpeechDirect: ~ \$7.5 K	~ \$120 K

Excluding device and network costs

<b>COST PER ADDITIONAL USER (List Price)</b>	Full Speech: ~ \$1,295 SpeechDirect: ~ \$550	~ \$1,750
--	---	-----------

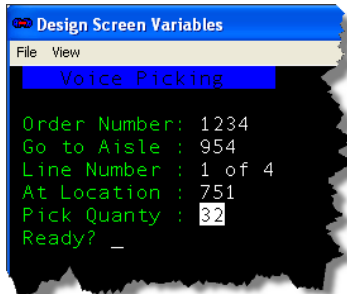
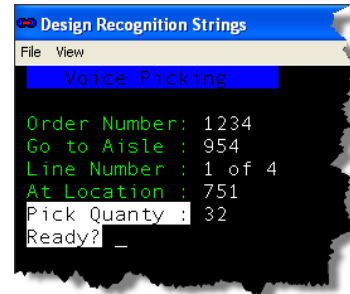
Excluding device and network costs

## SCREEN RECOGNITION & SPEECH RULES INTEGRATED INTO STAY-LINKED ADMINISTRATOR

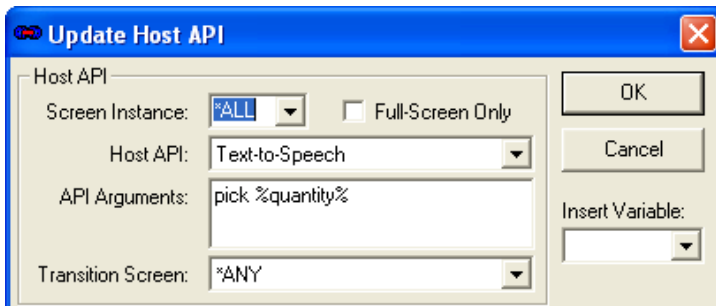
Screen ID	Enabled	Show	Cursor Row	Cursor Col	Screen Map
AS/400 Sign On Screen	True	False	0	0	AS/400 - Sign On Screen
AS/400 Sign On Screen w/CPF	True	True	0	0	AS/400 - CPF Messages
AS/400 Program Messages	True	False	0	0	AS/400 - Display Program Messages
AS/400 Main Menu	True	False	0	0	AS/400 - Main Menu
Voice Demo Pick Confirm	True	True	0	0	Voice Demo - Pick Qty Confirm
Voice Demo Location Check Digit	True	True	7	16	Voice Demo - Location Check Digit
Voice Demo Aisle Confirm	True	True	0	0	Voice Demo - Aisle Number Confirm
Voice Demo Order Confirm	True	True	0	0	Voice Demo - Order Number Confirm
Voice Demo Order Prompt	True	True	3	16	Voice Demo - Order Number Prompt
Voice Demo Splash Screen	True	False	0	0	Voice Demo - Splash Screen

Screens are captured and managed through the same Stay-Linked GUI Administrator used to control the terminal emulation environment. Can be run on any networked PC.

- Recognition strings – anything that is consistently shown on a particular screen – are defined.
- Supports multiple strings, as shown here – “Pick Qty :” in row 7, column 2 and “Ready” in row 8, column 2.
- Also supports the absence of other strings. (e.g., The screen has “Ready?” in row 8, column 2 but DOES NOT have “Error” in row 25, column 1.)



- Information fields on the screen can also be defined for use during the process.
- In this case, the “quantity” data will be saved every time the value changes.
- This provides the ability for the speech engine to tell the user to “pick 32” as this screen shows.
- The variable values are preserved between screens so they can be used on later screens – or even when returning to previous screens. (e.g., “you successfully pick 32 items from location 751.”)



Instructions for speech are defined using the Stay-Linked API dialog. In this case, the speech engine will announce to the user to “pick 32” when this screen displays. A second API is defined for this screen (not shown) instructing the engine to accept a quantity value.

## SOLUTION COMPARISON

(Stay-Linked SpeechDirect Excluded – No Comparison to Alternatives)

	Stay-Linked & CTGTALK	Vocollect	Voxware	Wavelink Speakeasy
<b>Architecture</b>	Host Terminal Emulation	Server Based	Server Based	Device Based
<b>Network Impact</b>	Low	High	High	High
<b>Reuses Existing Applications</b>	Yes	No	No	Yes
<b>Multi-modal Input from Device</b>	Always	Maybe	Sometimes	Yes
<b>Requires Speaker Training</b>	No	Yes	Yes	Yes
<b>Proprietary Headset</b>	No	Yes	Yes	?
<b>Speaker Independent Recognition</b>	Yes	No	No	Yes
<b>Easily Adaptable</b>	Yes	No	No	No
<b>Automatic Noise Filtering</b>	Yes	No	No	No
<b>Implementation</b>	5 – 10 Days	Months	Months	Weeks to Months
<b>Additional Server Required</b>	No	Yes	Yes	Yes
<b>Approach</b>	Software Only	Hardware	Software	Devices do all the work
<b>Advantages</b>	Leverage existing Host application logic by adding Speech on top	Provides a front-end warehouse application you have to interface	Provides a front-end warehouse application you have to interface	Leverage existing Host application logic by adding Speech on top
<b>Disadvantages</b>	None	Need to create interfaces back to your Host applications	Need to create interfaces back to your Host applications	All rules must be copied to each device ----- Proxy Server needed for reliability