

논문접수일 : 2012.10.05

심사일 : 2012.10.12

게재확정일 : 2012.10.27

# 물류관리시스템 환경에서 사용자 중심 음성인식 인터랙션 시스템 디자인을 위한 연구 프레임 및 방법 개발

Development of Analytic Research Frame and Methods for Enhancing Voice-driven  
Interaction in a Warehouse Management System

**최 정 민**

국립공주대학교 산업디자인공학부 시각정보디자인전공

**Choi, Jung-Min**

School of Industrial Design, National Kongju University

## 1. 서론

## 2. 이론적 배경

- 2.1 음성인식 인터랙션
- 2.2 물류관리시스템 유형
- 2.3 음성인식 물류관리시스템의 인터랙션

## 3. 음성인식 물류관리시스템 사례연구

- 3.1. Talkman 프로젝트 개요
- 3.2. 사용자 조사 및 데이터 수집

## 4. 음성인식 디자인 연구 프레임 구성

- 4.1. 거시적 분석(Macro-Level Analysis) 방법
  - 4.1.1 사용자(Users), 행동(Activities), 기기(Objects) 분석
  - 4.1.2 사용자, 행동, 기기 통합 분석
- 4.2. 미시적 분석(Micro-Level Analysis) 방법
  - 4.2.1 조작흐름 분석(Operational Flow Analysis)
  - 4.2.2 기기 상태전이 분석(State Transition Analysis)
  - 4.2.3 조작흐름, 기기 상태전이 통합 분석
  - 4.2.4 사용자 프로필 분석(User Profile Analysis)

## 5. 제안된 프레임을 적용한 사례연구 성과

- 5.1 거시적 분석 방법의 성과
- 5.2 미시적 분석 방법의 성과

## 6. 결론

## 참고문헌

## 논문요약

물류 관리시스템의 효과적 관리를 위한 정보시스템에 대한 연구의 필요성이 증가되고 있으며, 특히 음성인식을 기반으로 하는 휴대용 정보장치의 개발이 점차 활발해지고 있다. 본 연구의 목적은 음성인식기술을 보다 효과적으로 적용하기 위한 사용자 중심 음성인식 디자인을 위한 연

구 프래밍과 방법을 제안하는 것이다. 이를 위해, 이론 연구를 통해 음성인식 인터랙션의 개념과 사용자중심 디자인을 위한 기본 원칙을 도출하였다. 그리고, 이를 기반으로 물류관리시스템으로 활용되는 Vocollect사의 Talkman을 프로젝트를 사례로, 사용자 조사 결과를 사용자 중심의 관점에서 분석하기 위한 연구 프레임 및 방법을 구성하여, 향후 유사한 연구에 활용할 수 있도록 하였다. 제안된 프레임 및 방법의 기본 틀은 전체 환경 분석을 위한 AEIOU + Time을 결합한 프레임과 거시적, 미시적 분석을 위한 세부 방법과 분석을 위한 모형을 구성의 원리와 함께 제시하였다. 또한 이 모형을 사례 분석을 결과와 함께 제시하여, 향후 음성인식 기반 시스템과 제품을 개발하는데 응용될 수 있도록 하였다.

## 주제어

음성인식, 인터랙션 분석 방법, 음성인식 인터랙션 디자인

## Abstract

Effective convergence of mobile information technologies to package delivery services increasingly becomes important for delivery performance and better experience for both workers and customers. In order to improve the current experiences with hand-held devices, this paper examines the design opportunities for applying voice recognition technologies to mobile information devices used in package delivery service industry. By conducting observational user studies and analyzing the collected data from multiple viewpoints, the present research found some unmet needs in the current device usage, including needs for hands-free during moving and organizing parcels, ease of operations with wearing gloves, and seamless telecommunications with the office and customers. This research proposes that these kinds of issues will be able to be better addressed by adopting voice-driven interaction in package delivery processes. The future investigation will involve the development of voice-based information device design for package delivery and the evaluation of the effectiveness and experiences through user studies.

## Keywords

Voice Recognition, Analytic Methods for Interaction, Voice-driven Interaction Design

## 1. 서론

오늘날 테크놀러지의 발전과 함께 제품과 사용자 간 커뮤니케이션, 인터렉션 모달리티(Modality)는 점점 더 다양화 되고 있다. 과거 디지털 제품 및 시스템 디자인 초창기에는 주로 물리적 도구를 사용한 입력과 시각 피드백을 위주로 한 커뮤니케이션 방식이 주를 이루었지만, 제품들이 인간의 일상 생활에 스며들어 다양한 가능성을 확장시켜주게 되면서 좀 더 편리하고 즐거운 사용자 경험을 제공해주기 위해 다양한 대화 방식이 모색되어왔다. 특히 음성을 활용한 청각 모달리티(Auditory Modality)는 이미 오래 전부터 모바일 기기, 각종 가전제품 등 다양한 디지털 제품의 디자인에 많이 활용되어 왔다. 그리고 최근 유무선 통신 분야에서 음성인식(Voice-Recognition) 기술이 급속한 발전을 하고 있으며, 이제는 다양한 분야로 확산되고 있다. 음성인식 기술의 발전과 활용은 사용자-제품 인터랙션의 미래를 크게 바꿔놓을 것으로 기대되고 있다(김상훈, 이영직, 2004).

음성인식 활용 제품은 일반 소비제품 보다는 특수한 산업적 필요를 충족하는 방향에서 많은 시도와 성과가 있어왔다. 특히, 전자상거래의 발달로 인해 물류 관련 기반산업이 급속도로 성장하고 있으며, 이에 따라 음성인식 기반 기술의 물류관리시스템(Warehouse Management System)에서의 활용은 여러 가지 장점으로 인해 음성 인식 분야 중에서도 가장 빠른 성장세를 보이는 분야 중 하나이다(Raiyana & Kumar 2006).

무엇보다 다양한 물류를 다루어야 하는 상황에서 핸드프리가 가능하고, 빠르고 정확한 주문처리가 용이하며, 주문처리 과정의 다양한 사용자간의 효율적 커뮤니케이션이 가능한 등의 많은 장점이 있다. 또한 음성은 사람에게 가장 자연스럽게 편안한 정보 커뮤니케이션의 수단이라 할 수 있을 것이다. 하지만, 여전히 음성인식 인터랙션은 실제 사람들이 이야기하는 자연어와 음성인식 기기가 이해할 수 있는 언어를 어떻게 구분할 것인가를 비롯, 해결해야 할 많은 과제들을 안고 있다.

물류 시스템에서의 음성인식 기기들은 산업적 요구에서 출발한 특성상 시스템의 효율성과 생산성 향상 측면에서 주로 관심을 가져왔다. 즉, 보다 빠르고 정확한 오더 프로세싱 기술이 주된 관심사였으며, 사용자의 경험에 대해서는 관심이 적었다. 그러나 음성인식 분야의 연구는 이제 비용절감이라는 ROI 측면의 초점에서 사용자에게 만족을 줄 수 있도록 메시지를 전달하는 것으로 연구의 주요 관심이 옮겨지고 있

다. 물류관리시스템의 사용자 경험에 대한 관심은 음성인식 시스템의 사용성 향상에 직접적 영향을 미치며, 궁극적으로는 시스템 사용 수행도(performance)의 향상에도 영향을 미칠 수 있다.

이와 같이 물류 환경에서의 음성인식 적용에 대한 연구는 중요성이 커지고 있으며, 특히 사용자 경험에 대한 관심이 점차 증가하고 있다. 따라서 본 연구에서는 물류 환경에서 인간 중심의 음성인식 시스템 디자인을 위한 디자인연구 프레임 및 방법을 제안하는 것을 목적으로 하고 있다. 이를 위하여, Vocollect사(www.vocollect.com)에서 생산되는 음성인식 제품을 사용하는 물류회사를 대상으로 사용자 조사, 시장분석 등을 수행하여, 음성인식 시스템을 디자인하기 위한 연구 방법을 제안하고, 실제 사례를 들어 제안된 연구 방법을 평가하고자 한다.

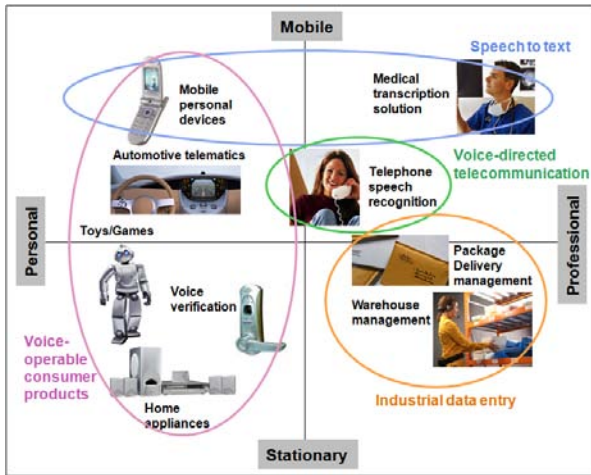
본 연구의 의의는 앞으로 많은 발전이 예상되는 음성인식 분야의 인터랙션 디자인의 기본 방법을 제시하여, 보다 사용자 중심의 디자인 연구를 위한 기초를 마련하는데 있다.

## 2. 이론적 배경

### 2.1 음성 인식 인터랙션

음성 인식 기술은 모든 사람들이 보편적으로 사용하는 말을 통해 시스템의 조작 및 상호작용을 가능하게 하는 기술이다. 음성인식기술은 새로운 입력방법으로 1960년대부터 연구가 시작되어, 1990년 후반부터 실제 시스템에 적용되기 시작하였다. 인터랙션 디자인에서 입력기술은 초기 키보드에서 마우스와 같은 보다 직관적인 방법으로 진화를 거듭하고 있으며, 직관적 조作的 터치 입력 방법과 함께 음성인식 기술은 새로운 대안으로 떠오르고 있다. 음성인식은 무엇보다도 인간의 언어를 활용한다는 점에서 인간에게 만족감을 줄 수 있는 인터페이스 방법으로 떠오르고 있다.

‘음성인식 기술의 개요와 최근동향’ 보고서에 따르면, 음성인식 활용 분야는 매우 다양하여, 산업 전반에 걸쳐 활용되고 있다. 특히, 지식 베이스(knowledge base)와 기술의 결합을 요하는 첨단 기술로써 컴퓨터, 휴대폰, 녹음기, 전자사전, 게임기, 장난감, 멀티미디어 가판대, 가정용 전자기기 등의 제품에 채택돼 음성을 통한 다이얼링, 정보검색, 회의록 작성, 게임, 학습, 전자제품 제어, 보안 설정 등을 가능하게 한다. 최근 음성인식기술은 미국을 중심으로 구체적인 응용분야가 개척되고 있고 멀티모드/멀티미



[그림 1] 다양한 음성인식 기술의 활용 현황

디어 환경 속에서의 다른 미디어와 통합에 관한 연구들이 진행되고 있다.

[그림 1]은 음성인식 기술이 활용되는 다양한 영역을 보여주고 있다. 개인적으로 활용되는 분야는 휴대폰, 자동차의 음성 조작, 로봇이나 게임과 같은 영역에서 많이 활용되고 있다. 또한 기업의 콜 서비스 센터 등의 고객관리 시스템도 음성인식 기술이 많이 활용되는 분야이다. 하지만, 이중에서도 전문 산업적 영역에서 가장 많이 활용되고, 현재 가장 성공적인 영역은 물류 및 택배와 같은 배달서비스 비즈니스로 평가되고 있다.

## 2.2 물류관리시스템 유형

물류관리시스템(Warehouse Management System, WMS)의 입출력 방식은 크게 페이퍼 방식, RFID(전자식별태그) 방식, 그리고 음성인식 방식의 세 가지로 구분되어진다.

페이퍼 방식은 모든 주문에서부터 주문된 물품의 선택 및 배송의 과정이 종이로 이루어지는 가장 오래된 방식의 물류관리시스템이다. 페이퍼 방식의 경우, 전체 주문 목록을 한눈에 확인하거나, 한번 물류창고 내의 물건의 위치에 익숙해지면 빠르게 과업을 수행할 수 있는 장점이 있지만, 주문한 물건이 없거나 새로운 물건이 들어오는 경우 일일이 처리해야하는 불편함이 존재한다. RFID 방식의 경우, 물건이 보관되는 선반에 각 물건을 구분할 수 있는 RFID를 설치하여 운영하는 방식이다. 물건의 재고나 시스템화하여 입출고나 재고 관리 등을 할 수 있는 장점이 존재하지만, 실제 주문을 처리하는 과정에서의 효율성을 증진시키기는 쉽지 않다. 마지막으로 음성 주도(Voice-driven) 방식의 물류관리시스템이 있다. 이 방

식은 물류관리시스템에서 가장 노동력 집약적인 부분인, 주문한 물건을 모아서 포장하는 오더피킹(Order Picking)에 가장 장점이 있는 방식이다. 오더피킹 작업은 그 특성상 손이 자유로워야 하고 빠르고 정확한 작업이 이루어져야 하므로, 음성주도 기술이 최적의 솔루션이라는 평가를 받고 있다.

## 2.3 음성인식 물류관리시스템의 인터랙션

물류 관리의 효율성을 높이고자 음성 주도의 시스템이 도입되었지만, 사용자의 관점을 기반으로 이루어진 연구는 많지 않다. 음성인식이 기술주도적 성격을 가지기도 하였지만, 아직 시장에 도입된 초기단계를 막 지난 시점이기도 하여, 아직은 음성인식 시스템의 기술적 성능에 더 많은 관심이 있기 때문이다. 하지만, 음성인식 시스템의 경우, 음성을 매개로 여러 그룹의 사용자들이 유기적으로 함께 일을 하는 것이므로, 사용자의 경험이 일의 생산성에도 영향을 미칠 것이며, 따라서 사용자 관점에서의 음성기반 물류시스템을 연구하는 것은 매우 중요하다고 할 수 있다

음성인식 시스템이 다른 인터페이스 방식에 비해 차별화되고 유리한 방식임을 주장하는 많은 연구가 있다. 무엇보다, 직관적이고 효율적인 상호작용이 가능하고, 항상 휴대가 가능한 유비쿼터스적인 특성을 가지고 있으며, 즐거운 사용경험을 제공하고, 손을 자유롭게 쓸 수 있고, 상호작용 상황에서 시선을 빼앗기지 않아도 되는 장점이 있다(Cohen et al., 2004). 하지만, 이에 대한 반론도 존재하는데, 시각정보를 제공받는 경우보다 실제 인지적 부담이 더 커서 오히려 효율적이지 않다는 의견도 있다.

여러 가지 이견에도 불구하고, 음성기반의 사용자 경험을 향상시키기 위한 연구는 매우 필요한 시점이며, 본 연구는 물류시스템을 사례로 하여 음성기반의 사용자 인터랙션 분석을 위한 기본 방법을 제안하고자 한다.

## 3. 음성인식 물류관리시스템 사례연구

### 3.1. Talkman 프로젝트 개요

본 논문의 목적인 물류환경에서 사용자 중심 음성인식기반 시스템의 연구방법 제안을 위해, 현재 사용되고 있는 음성인식 물류관리시스템을 사례로 조사하였다. 사례연구 프로젝트는 미국 물류시장에서 가장 많이 사용되는 VoCollect 회사의 음성인식 기반의 휴대용 물류관리시스템을 대상으로 하였다. [그림 2]는



[그림 2] 물류창고 환경(좌측 상단), Talkman(우측 상단), 헤드셋

본 연구에서 관찰한 물류창고의 작업 환경, 개인 작업자들이 휴대하는 터미널인 Talkman, 그리고 터미널과 함께 사용되는 헤드셋을 보여주고 있다.

### 3.2. 사용자 조사 및 데이터 수집

본 연구를 위해 미국 시카고 북쪽에 위치한 물류도매창고에서 사례조사를 진행하였다. 일반적으로 각

소매업체 또는 편의점들이 필요한 물건을 대량으로 물류도매창고에 주문하여 물건을 공급받아 일반 소비자에게 판매를 하고 있다. 이러한 물류도매창고는 음성인식기반의 인터랙션 기기를 활용하여, 주문접수, 주문받은 물건을 물류창고에서 수거하고 이를 패키지화하여 해당되는 트럭에 실을 수 있도록 배달하는 역할을 수행하고 있다. 물류도매창고 작업에는 다양한 사용자 그룹이 관여되어 있다. 소매상(Retailer)을 통해 주문을 접수 받고 전체 프로세스를 관리하는 관리자(Manager), 실제 주문을 처리하는 주문처리자(Selector), 그리고 부족한 물건을 보충하는 재고관리자(Forklift Operator)가 있다. 본 연구에서는 음성인식시스템의 사용경험 평가를 위해, 각각 세명의 숙련된 사용자와 초보사용자 총 6명 대상으로 사용경험이 어떻게 다른지를 분석하였다. 분석을 위해, 6명의 사용자를 관찰하고 후속인터뷰를 진행하였다. 관찰은 주문이 접수되었을 때 해당 주문을 처리하는 전체 과정을 비디오로 녹화하였으며, 픽업 과정에서 특히 행동을 보인 이유 혹은 현재 사용되고 있는 시스템에 대한 의견을 사후 인터뷰를 통해 수집하였다.

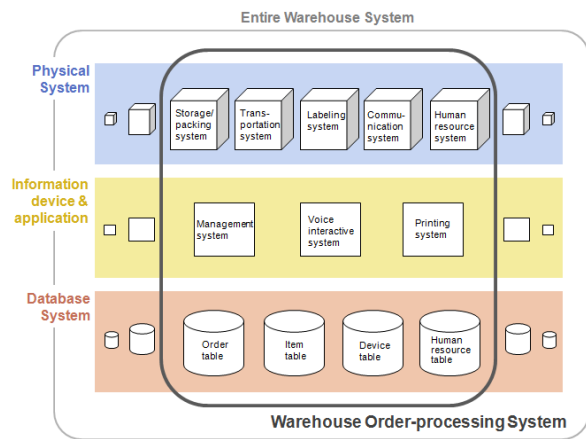
	Activities	Users	Objects	Environment	Interaction
Before selecting	Enter order into system	Retailer		Customer's retail store	Between DB & outside systems (data transmission)
	Transfer order data to manager				
	Receive orders from retailers	Manager	Database system	Operational office	
	Confirm the order		Order lists from retailers		
	Receive products from suppliers		Wireless network equip.		
	Manage database		Wired network equip.		
	Queue orders		Queued list		
	Manage voice system		Voice system manager		
	Drive forklift	Forklift operator	Forklift	Shelving area	Between forklift operator & forklift (physical interaction)
	Replenish		Boxes		
Create voice template	Selector		Operational office		
Adjust voice properties					
While selecting	Sign in Talkman		Talkman terminal		Between selector & Talkman (electronic voice interaction)
	Receive summary info. of order		Headset / Microphone		Between Talkman & DB (data transmission)
	Write down summary of order		Belt		
	Drive pallet truck		Cord & alligator clip	Shelving area	Between selector & truck/boxes (physical interaction)
	Receive location info. of product		Pallet Truck		Between selector & numbering system (visual interaction)
	Receive quantity info.		Pallets		
	Confirm quantity		Shelves		
	Read check digits		Aisle numbers		
	Pick up boxes		Slot numbers		
	Organize boxes on pallets		Check digits		
	Make broadcasting communications	Manager	Memo paper	Operational office	Between selectors & manager (personal interaction via speakers)
	Ask Talkman about current status	Selectors	Reacher (skewer)	Shelving area	
	Talk with colleagues		Speakers in warehouse		Between selectors (face-to-face interaction)
Report shortage		Fans			
Generate chase order		Chase order sheet			
After selecting	Wrap each pallet		Shrink-wrap	Printing area	Between selector & labels (physical & visual interaction)
	Order label-printing		Printer		
	Affix labels onto wrapped boxes		Printed labels to identify order		
	Drop order off at docking deck		Bay numbers	Docking deck area (bay)	
	Sign off Talkman		Battery charger		
	Receive performance statistics				Between selector & charger (physical interaction)
	Charge battery			Operational office	Between truck driver & truck (physical interaction)
	Assign order to truck	Manager			Between truck driver & retailer (face-to-face interaction)
Load/unload items	Truck driver	Truck	Docking deck area (bay)		
Transport items					
Receive delivery document	Retailer	Delivery document	Customer's retail store		

[그림 3] AEIOU + Time 프레임



#### 4. 음성인식 디자인 연구 프레임 구성

물류창고에서의 음성인식 시스템 연구를 위해서는 두 가지 차원, 즉 거시적 분석(Macro-Level Analysis)과 미시적 차원의 분석(Micro-Level Analysis)이 필요하다. 전체 물류창고 업무의 맥락을 이해하는 것, 즉, 어떤 사용자 유형이 존재하며, 각 사용자의 행동 양식, 그리고 어떠한 기기들이 활용되는가를 이해하는 것이 거시적 차원의 분석(Macro-Level Analysis)에 해당된다. 그리고 실제 물건을 수집하는 Selector가 어떤 방식으로 작업하는지, 주문을 처리하는 과정에서 다른 사용자들과의 커뮤니케이션과 상호작용은 어떻게 이루어지는지 면밀히 분석하는 것이 미시적 차원의 분석(Micro-Level Analysis)에 해당된다. 이를 위하여, Doblin그룹의 Rick Robinson과 Stef Norvaisis에 의해 제안된 분석 방법인 AEIOU 프레임워크를 사용하되, [그림 3]에서 보는 바와 같이 시간의 개념을 추가하여, 분석 대상을 보다 총체적으로 설명할 수 있는 프레임을 제시하였다. 즉, X축은 사용자 행동 분석을 위한 다섯 가지 요인 - Activities, Environment, Interactions, Objects, Users - 을 사용하였으며, Y축은 실제 Selector의 업무를 시간 순으로 배치하여, 음성인식 시스템이 사용되는 물류창고 업무의 전체 맥락을 설명하는 틀로 제시하였다.



[그림 4] Modeling of system structure

[그림 4]는 관찰과 인터뷰를 통해 파악한 해당 물류창고시스템의 주문관리 정보 구조를 보여주고 있다. 실제 음성인식 시스템을 모형화하여 인터랙션 디자인을 제안하기 위해서는, 전체 물류시스템을 분석하고 음성인식 시스템이 적용될 수 있는 가능성과 적절한 적용의 방법은 분석하는 것이 중요하다. 이러한 차원에서 거시적 분석은 그 의미가 있다고 할 수 있다.

#### 4.1. 거시적 분석(Macro-Level Analysis) 방법

##### 4.1.1 사용자(Users), 행동(Activities), 기기(Objects) 분석

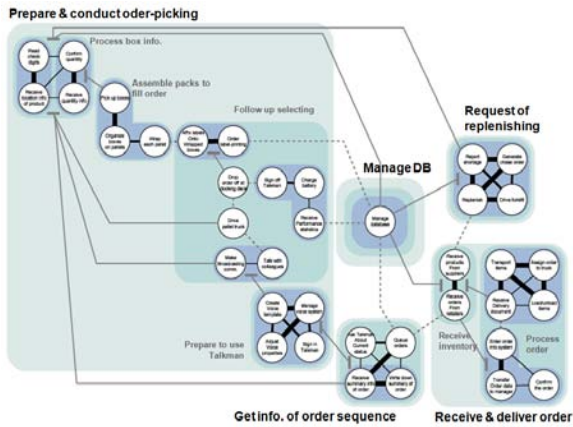
물류창고에서 이루어지는 사용자의 행동 분석을 위해, 관찰된 행동 간의 관계성을 [그림 5]와 같이 매트릭스로 분석하였다. Insight Matrix라고 명명된 이러한 형식의 프레임은 Kumar(2008)에 의해 제안된 것으로, 각 요소 간의 유사성을 기반으로 점수를 부여, 컴퓨팅 처리에 의해 개별 요소들을 군집으로 묶음으로써 정보의 패턴을 읽는 방식이다.

	Receive location info of product	Read check digits	Confirm quantity	Receive quantity info	Pick up boxes	Organize boxes on pallets	Wrap each pallet	Affix labels onto wrapped boxes	Order label-printing	Drop order off at docking deck	Drive pallet truck	Talk with colleagues	Make broadcasting comm.	Identify box	Process box info.	Assemble	Assemble	Prepare, load, & follow up selecting
Receive location info of product	3	3	0	0	1	0	0	0	0	0	0	0	3	1	1	0	0	0
Read check digits	3	3	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0
Confirm quantity	0	1	3	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Receive quantity info	0	1	3	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Pick up boxes	1	1	2	2	3	3	0	0	0	0	0	0	0	0	0	0	0	0
Organize boxes on pallets	0	0	0	0	3	3	2	0	0	0	0	0	0	0	0	0	0	0
Wrap each pallet	0	0	0	0	2	3	1	0	0	0	0	0	0	0	0	0	0	0
Affix labels onto wrapped boxes	0	0	0	0	0	1	3	3	1	0	0	0	0	0	0	0	0	0
Order label-printing	0	0	0	0	0	0	3	3	1	0	0	0	0	0	0	0	0	0
Drop order off at docking deck	0	0	0	0	0	0	0	1	1	3	2	0	0	0	0	0	0	0
Drive pallet truck	3	0	0	0	0	0	0	0	0	2	3	1	0	0	0	0	0	0
Talk with colleagues	1	1	0	1	0	0	0	0	0	0	1	3	0	0	0	0	0	0
Make broadcasting comm.	1	1	0	1	0	0	0	0	0	0	0	0	3	0	0	0	0	0

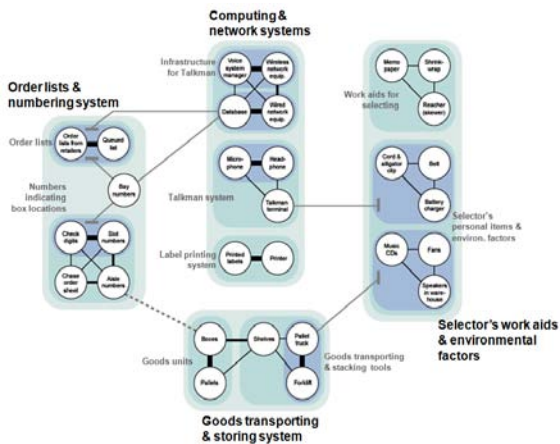
[그림 5] 사용자 행동 분석 매트릭스(부분 확대)

Insight Matrix에서 도출한 물류창고 업무(행동) 군집 간의 관계를 보다 직관적으로 파악하기 위해서는 [그림 6]과 같이 네트워크 다이어그램의 형식으로 표현할 수 있다. 이 분석을 통해, 물류창고의 업무는 주문 데이터 관리를 중심으로 이루어지며 이는 음성 시스템에 의해 구성되는 것을 알 수 있다.

한편, 사용자 행동 분석에 사용된 것과 동일한 방식을 적용, 사용되는 기기들 간의 관계를 분석하여 [그림 7]과 같은 네트워크 다이어그램으로 표현하였다.



[그림 6] 사용자 행동의 네트워크 다이어그램



[그림 7] 사용 기기의 네트워크 다이어그램

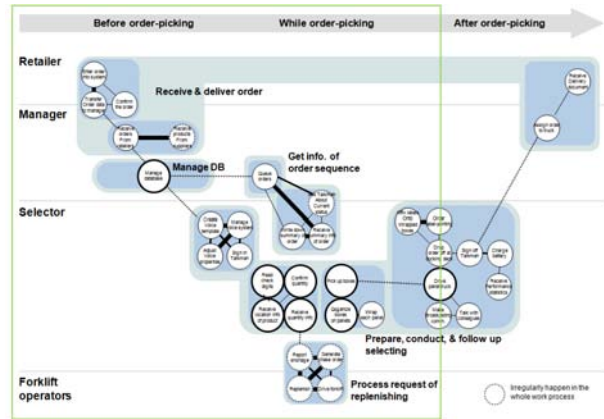
이와 같은 사용자, 행동, 기기의 개별 분석을 통하여, 연구자는 각 사용자 그룹별 업무와 업무 특성을 이해할 수 있다. 행동 분석을 통하여, 물류창고에서의 업무는 크게 다섯 가지로 구분되며, 특히 음성인식 시스템이 주문을 관리, 처리, 배송하는 시스템과 긴밀하게 연동되는 것이 중요하다는 것을 알 수 있다.

#### 4.1.2 사용자, 행동, 기기 통합 분석

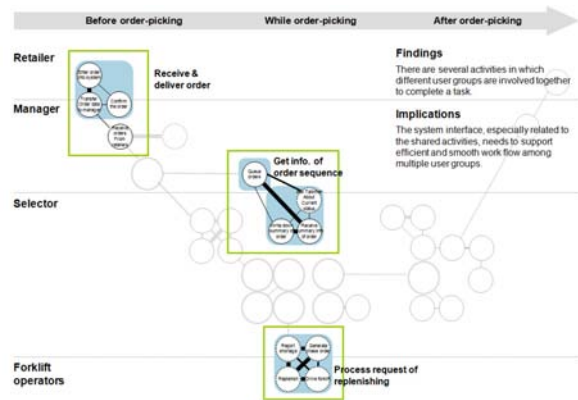
개별적으로 수집된 사용자, 행동, 기기를 통합하여 분석하는 것은, 음성인식 기반 물류창고 환경을 이해하기 위해 매우 중요한 방법이다.

우선 첫 번째 방법은 [그림 6]의 분석 결과를 X축을 시간으로, 그리고 Y축을 사용자로 구성된 분석틀에 재구성하는 방법이다. 즉, 전체 행동은 주문처리 전, 중, 후로 구분이 가능하며, 각 행동은 그 행동을 수행하는 사용자가 있다. 이러한 점을 고려하여 새로운 분석 방법을 [그림 8]과 같이 제안하였다.

[그림 8]과 같은 통합적으로 모형화를 하게 되면, 음성인식을 기반의 시스템 디자인에서 중요한 시사점을 얻을 수 있다. [그림 9]는 이러한 시사점을 잘 보



[그림 8] 분석된 행동을 업무순서, 사용자 분석 결과와 통합하여 분석하기 위한 프레임



[그림 9] 음성인식 디자인 기회 분석 프레임

여주고 있는데, 음성인식 시스템 디자인에 있어서, 각 유저간의 접점에서 디자인 문제가 많이 발생함을 한 눈에 알 수 있다. 소매상(Retailer)과 관리자(Manager) 사이에서 생성된 주문을 확인하고 접수하는 과정이 매우 중요함을 알 수 있었으며, 또한 관리자(Manager)와 주문처리자(Selector)가 주문의 내용을 음성인식을 통해 정확히 전달/확인하여 처리하는 과정도 중요하다. 특히, 청각 모달리티 인터랙션의 특성과 관련한 중요한 시사점을 발견할 수 있었는데, 주문처리자는 패키지를 효율적으로 쌓아나가기 위해 남은 패키지의 분량을 계속 확인하고자 하는 경향이 있었다. 음성인식 시스템을 활용하는 경우 적절하게 패키지를 구성하기 위해 전체 주문량을 직관적으로 확인하기 어렵다. 따라서 주문처리자가 지속적으로 '주문 목록이 얼마나 많이 남았는가?'를 계속 확인하게 되는데, 이러한 전체 주문량을 확인할 수 있게 해주는 UI(User Interface) 디자인이 매우 중요함을 파악할 수 있었다. 또한 재고가 없을 경우, 재고를 파악하여 선반에 물품을 재충전하는 과정을 위한 음성인식 UI디자인도 매우 필요한 것임을 알 수 있었다.

## 4.2. 미시적 분석(Micro-Level Analysis) 방법

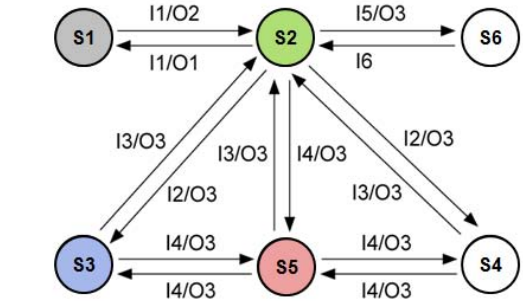
### 4.2.1 조작흐름 분석 (Operational Flow Analysis)

거시적 분석을 결과를 바탕으로, 실제 음성인식 시스템 인터랙션 분석을 위하여 [그림 10]과 같이 미시적 관점에서 사용자와 시스템의 인터랙션을 모형화하였다. 특히, [그림 9]를 통해 밝혀진, 인터랙션의 문제에 집중하여 모형을 작성하였다. 이러한 분석 프레임워크를 통해 밝혀 낼 수 있는 것은, 음성인식 시스템의 대화 방식을 인간 중심적인 개념과 시스템의 효율성의 관점으로 어떻게 조화시켜 디자인 할 것인가를 고려할 수 있다는 점이다. [그림 10]의 분석 결과, 음성인식 시스템을 통한 인터랙션이 시스템의 효율적 커뮤니케이션을 고려하여 매우 단순한 표현으로 디자인 되었다는 점을 알 수 있었으며, 따라서 일상적인 대화와 큰 차이가 있기에 사용자는 일정한 학습의 시간을 필요로 하고 있다. 그리고 전체 업무 과정을 이해할 수 있는 UI에 대한 제공이 상대적으로 부족하여, 사용자가 업무를 위해 필요한 시간이나 업무의 총량을 판단하기 어려운 문제점을 가지고 있다는 것을 알 수 있었다.

### 4.2.2 기기 상태전이 분석(State Transition Analysis)

음성인식을 위한 인터랙션에서 같이 고려되어야 할 내용은 인터랙션 기기(Object)에 대한 분석이다.

음성인식을 위한 기기가 어떻게 물리적으로 조작되며, 이러한 물리적 조작의 결과로 음성인식 인터랙션을 어떻게 지원하는가를 상태 전이 분석을 통해 [그림 11]과 같이 모형화하여 설명 할 수 있다.



#### States

- S1 Off
- S2 Initial state (sign-off)
- S3 Order picking
- S4 Call investigation
- S5 Adjusting voice properties
- S6 Making new template



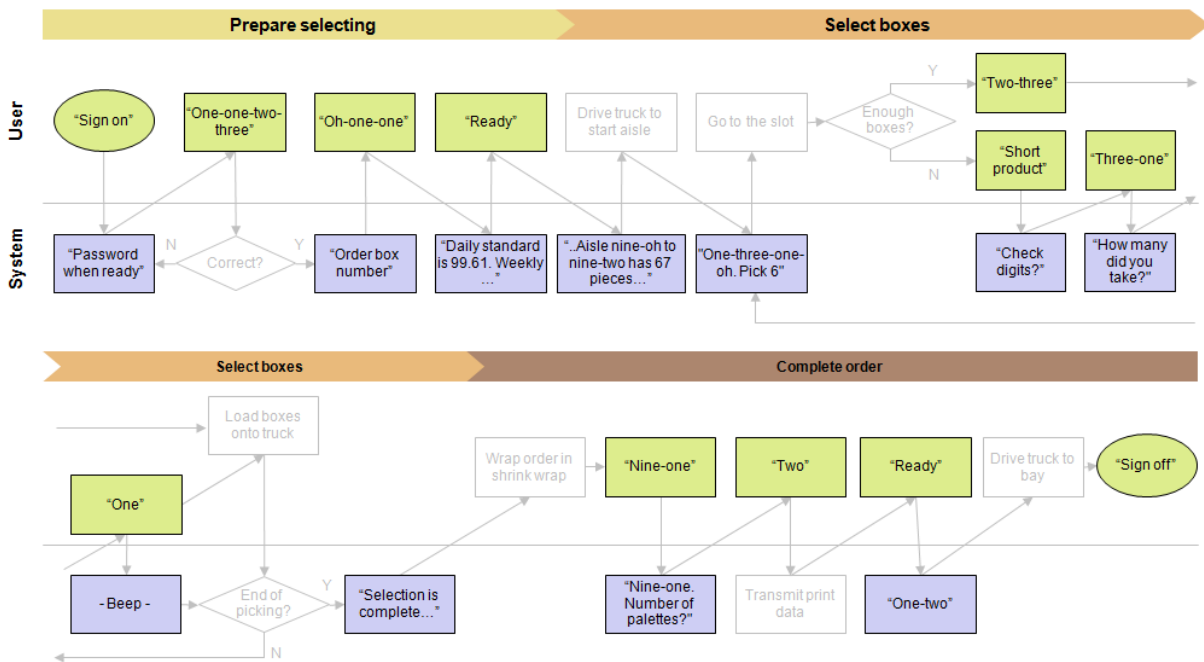
#### Input

- I1 Press Play/Pause button
- I2 Select menu by speech
- I3 Order sign-off by speech
- I4 Press Operator button
- I5 Connect visual training device
- I6 Disconnect visual training device

#### Output

- O1 Power Off
- O2 Flash green LED
- O3 Provide instruction by speech

[그림 11] 기기 상태전이 분석을 통한 음성인식 디자인 기획 분석 프레임



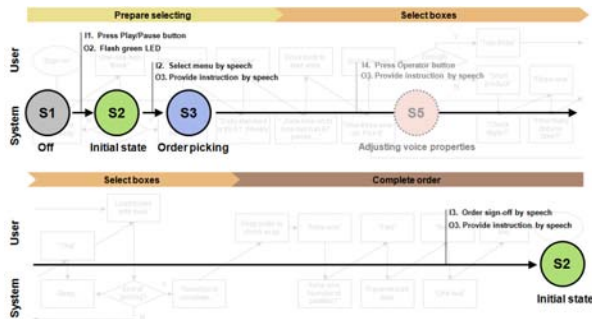
[그림 10] 음성 인식 UI에 기반한 조작 흐름 분석 (Operational Flow Analysis)



### 4.2.3 조작흐름, 기기 상태전이 통합 분석

앞서 거시적 분석에서 인터랙션 디자인 기회를 포착하기 위해 통합적 분석을 시도한 것처럼, 미시적 분석에서도 보다 통찰력있는 해석을 위해 [그림 10]과 [그림 11]을 통합하여 보다 다양한 해석을 이끌어 낼 수 있다.

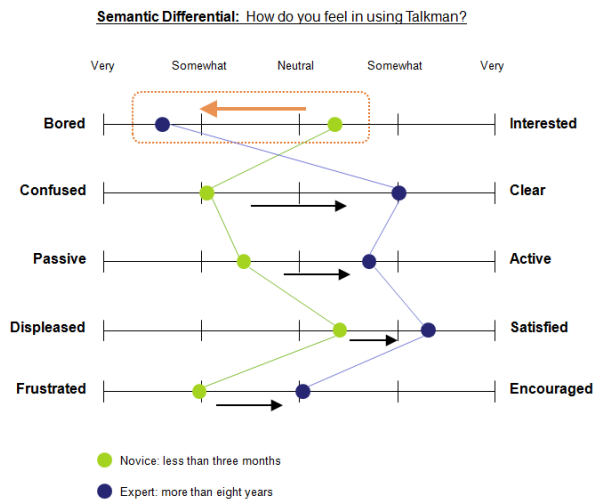
[그림 12]에서 보는 것처럼, 주문 처리를 위하여 S1, S2, S3가 가장 많이 활용되는 상태임을 알 수 있으며, 주문처리 중 음성인식과 관련된 설정 변경을 위해 S5 상태로 전이하는 경우가 많은데, 이럴 경우 하드웨어 버튼을 활용하는 것이 다소 불합리함을 직관적으로 알 수 있었다. 즉, 주문 처리를 위해서는 양손을 모두 사용하는 경우가 대부분인데, 갑자기 버튼 조작을 위해 다른 모달리티를 활용하는 것은 비효율적이라고 할 수 있다.



[그림 12] 조작흐름과 상태전이의 통합적 분석 프레임

### 4.2.4 사용자 프로파일 분석(User Profile Analysis)

조작흐름분석을 통해 설명된 것처럼, 현재 많은 음성인식 시스템은 적절한 학습을 필요로 한다. [그림 13]에서 보듯이 초기 학습 기간이 지나고 숙련도가 높아질수록 한 가지 단점이 나타나는데, 음성인식 시스템을 사용하는 것이 매우 단조롭고 지루한 일이 되어 가는 것이다. 이러한 것은 장기적으로 좋지 않은



[그림 13] 사용자 프로파일 분석 모형

업무 성과로 이어지는데 이러한 부분에 대한 개선점을 찾는 것이 향후 음성인식기반 시스템을 디자인하는데 중요한 시사점의 하나가 될 것으로 생각된다.

## 5. 제안된 프레임워크를 적용한 사례연구 성과

본 연구에서 제안한 프레임워크와 방법을 활용하여 VoCollect 사의 Talkman 제품의 물류창고시스템에서의 활용 사례를 분석한 결과는 전체적으로 음성인식시스템을 보다 인간중심적으로 개선할 수 있는 가능성을 여러 가지 차원으로 발견할 수 있었다는 점이다.

### 5.1 거시적 분석 방법의 성과

1) 주문처리자의 업무 과정에 따른 행동 분석을 통해, 주문처리자는 전체 주문의 양과 패키지 구성을 효율성을 위해 전체 주문 목록을 판단할 수 있는 음성인식 인터랙션을 기대하고 있었다. 즉, 배달을 위한 패키지 구성에서 보다 무겁거나 혹은 크기가 큰 물건을 아래에 적재하고, 가볍거나 작은 물건을 위에 적재하면 보다 쉽게 주문을 처리할 수 있다. 또한 재고의 유무, 동선의 고려 등도 매우 중요한 요소로 분석되었다. 하지만, 현재의 음성인식시스템은 개별 명령에 대한 처리에만 초점이 맞추어져 있다, 이러한 거시적 관점에서 사용자의 요구를 충분히 반영하고 있다고 보기 어렵다.

2) 또한, 물류창고내의 다양한 사용자와의 보다 효율적 협업을 위한 인터페이스의 구성이 요구된다. 즉 관리자가 주문을 주문 처리자에게 넘기고, 재고관리자도 필요한 재고의 양을 미리 확인하여 업무를 처리한다면 업무의 효율성과 각 사용자간의 원만한 협업체계가 잘 구성될 것이다. 하지만, 현재 이러한 요구를 반영할 수 있도록 음성인식시스템이 구축되지 않아 사후처리식으로 업무가 이루어지고 있다는 것을 파악할 수 있었다.

### 5.2 미시적 분석 방법의 성과

1) 음성인식시스템을 도입하는 것은 모달리티를 다양하게 하여, 인지적 부담을 줄이는 것이 가장 중요한 목표이지만, 분석을 통하여 나온 결과는 시스템 설계의 효율성을 고려하여 명령어를 지나치게 단순하고 제한된 용어로 활용하는 것이 오히려 인지적 부담을 높이는 것으로 나타났다.

2) 조작흐름 및 주문처리자의 명령어 사용 패턴을 분석한 결과, 주문처리자는 전체 업무의 로드맵을 궁극해 하였다. 이는 거시적 분석에서도 나타났는데, 음

성인식 인터랙션 방법이 페이퍼 기반의 주문처리 방식이 가지는 장점, 즉 전체 그림을 볼 수 있는 기능을 고려하지 않았기 때문으로 이해될 수 있다. 이러한 인터랙션에 대한 고려는 매우 중요한 시사점으로 해석될 수 있다.

3) 지나치게 구조화된 음성인식방식은 업무 자체를 지루하게 인지하여, 향후 오히려 역기능으로 작용할 수 있다. 이는 음성인식 시스템을 도입한 다른 시스템에서도 흔히 볼 수 있는 현상인데, 이러한 부분에 대한 고려가 필요할 것이다.

4) 음성인식시스템이 현재는 기계와의 대화에만 초점이 맞추어져 있으나, 다른 사용자와의 대화를 위한 2명 이상의 사용자가 참여할 수 있는 구조로 변경되는 것이 필요한 것으로 분석되었다. 따라서, 대화의 패턴을 다각적으로 분석하고 이를 반영할 수 있는 인터랙션에 대한 연구가 필요할 것으로 판단된다.

## 6. 결론

본 연구는 물류관리시스템에서 사용자 중심의 음성인식 시스템 제안을 위한 디자인연구 프레임 및 방법을 제안하는 것을 목적으로 진행하였다. 본 연구에서 사용자 중심의 음성인식 프레임 및 방법 구성을 위한 원칙은 크게 사용자 중심 인터랙션 관점에서의 원칙과 방법론적 관점에서의 원칙으로 요약될 수 있다.

- 사용자 관점의 원칙: 음성인식 인터랙션과 관련된 환경과 환경을 구성하는 요소를 고려해야 한다. / Multi-Modality의 장점을 활용하여 사용자의 인지적 부담을 줄일 수 있어야 한다.

- 방법론 관점의 원칙: 디자인 문제를 쉽게 분석할 수 있는 방법이 제시되어야 한다. / 다양한 요소를 통합적으로 분석할 수 있는 방법이 제시되어야 한다.

이러한 원칙을 기반으로 본 연구에서 제안한 프레임의 구성 및 방법을 단계별로 요약하면 아래와 같다.

### 프레임 1: AEIOU + Time

기본 원칙은 전체 환경을 분석하기 위한 요소의 추출과 이것을 모형화 하는 것이 중요하며, 음성인식 인터랙션의 특성을 고려, 사용자, 행동, 기기와의 인터랙션에 초점을 맞추어야 하며, 인터랙션에 시간 개념이 들어가 있기 때문에, 이를 가장 잘 표현할 수 있는 방법으로 구성하였다.

### 프레임 2: 거시적 분석 프레임

거시적 프레임에 사용되는 방법은 사용자, 행동, 기기를 분석하는 개별 방법들을 제안하였으며, 특히

본 연구에서는 사용자 + 행동, 사용자 + 기기, 행동 + 기기를 통합적으로 분석할 수 있는 방법을 제안하여, 분석 이해와 분석의 가능성을 높였다.

### 프레임 3: 미시적 분석 프레임

미시적 프레임에 사용되는 방법은 사용자, 행동, 기기의 세부적 내용을 분석하는 방법을 음성인식 인터랙션이라는 분야에 맞게 응용하여 제시하였다. 조작흐름 분석에서는 음성을 매개로 사용자와 기기의 인터랙션에 초점을 맞추었다. 미시적 분석에서도 통합적 분석의 방법을 제시하여, 보다 다각적 관점에서 디자인 문제를 해석하는 것을 가능하게 하였다. 이러한 통합적 분석을 위한 방법은, 인터랙션에 관련된 기본 요소들의 공유점을 찾아내고, 그 공유점을 기준으로 요소들을 의미있게 재배열하는 것이 매우 중요하다.

본 연구에서 제시된 연구 프레임 및 방법은, 앞으로 많은 발전이 예상되는 음성인식 분야의 인터랙션 디자인의 연구의 기본 방향을 제시하여, 사용자 중심 음성인식 인터랙션 디자인 연구의 기초가 될 것으로 기대한다.

## 참고문헌

- 김남수 (1999). 「음성인식 기술의 개요와 최근 동향」. 뉴미디어 통신 공동 연구소 소식지.
- 김상훈, 이영직 (2004). 「신성장동력산업의 기반, 음성인터페이스 기술」. ETRI CEO information, 12.
- 손미숙, 박준석, 한동원, 조일연 (2006). 「웨어러블 시스템 사용자 상호작용 시장 분석 및 기술 동향」. ETRI 전자통신동향분석, 21(2).
- Aucella A., Kinkead, R., & Wichansky, A. (1987). Voice: Technology Searching for Communication Needs. *Proceedings of SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface*, ACM Press, 41-44.
- Cohen, M. H., Giangola, J. P., & Balogh, J. (2004). *Voice User Interface Design*. Boston: Addison-Wesley.
- Kumar, V. (2008). Design Strategies Reshaping the Innovation Landscape. Presentation at 33<sup>rd</sup> Annual Design Management Conference.
- Raiyani, S. & Kumar, J. M. (2006). Multimodal Warehouse Application. *Interactions*, 13.
- Suhm, B., Myers, B., & Waibel, A. (2001). Multimodal Error Correction for Speech User Interfaces. *ACM Transactions on Computer-Human Interaction*, 8(1), 60-98.