"Nobody Speaks that Fast!"
An Empirical Study of Speech Rate in Conversational Agents for People with Vision Impairments

논문 리뷰

Category

1. Questions
2. Method
3. DISCUSSION AND DESIGN IMPLICATIONS
4. Limitations & Future work

1. Questions

Q1. 시각장애인들이 정말로 CA가 빨리 말하기를 원하는가?

Q2. CA의 말하기 속도가 User Experience 혹은 Usage Patterns에 어떠한 영향을 미칠 것인가?

cf) CA(Conversational Agents) : Voice User Interface를 고려하여, 시각장애자들에게 새로운 기회를 제공하는 도구

2.1. Method : Participants

- 모바일 디바이스의 스크린 리더를 사용한 경험이 있는 시각장애인들
- 나이 : 27~58세 (평균나이 43.7세 표준편차 11.6)

cf) 해뜰폰 : 시각장애인용 맞춤 스크린리더 폰

ID	Age	Gender	Self-reported vision	Household size	Mobile	Speech rate of screen reader they used	Mobile CA they have experienced	Stand-alone CA they have experienced
1	58	M	Total blindness (Later in life)	Family (4)	Haetteul-phone	5.12 SPS	None	None
2	45	F	Total blindness (Later in life)	Alone	Haetteul-phone	7.23 SPS	None	SKT NuGu*
3	57	M	Total blindness (Later in life)	Family (2)	Haetteul-phone	6.83 SPS	None	None
4	54	F	Total blindness (From birth)	Family (2)	Haetteul-phone	5.34 SPS	Samsung Bixby	None
5	27	F	Total blindness (From birth)	Family (3)	iphone	9.13 SPS	Apple Siri	None
6	52	F	Total blindness (From birth)	Family (2, husband with low vision)	Haetteul-phone	6.65 SPS	None	None
7	37	F	Total blindness (From birth)	Family (4, husband with low vision)	Galaxy note	8.23 SPS	Samsung Bixby	None
8	45	M	Total blindness (Later in life)	Family (2, wife with blindness)	Haetteul-phone	6.01 SPS	None	None
9	29	M	Total blindness (From birth)	Alone	iphone	11.23 SPS	Apple Siri	Kakao mini**, KT Giene**
10	33	M	Low vision (Later in life)	Alone	Galaxys	5.56 SPS	None	SKT NuGu**

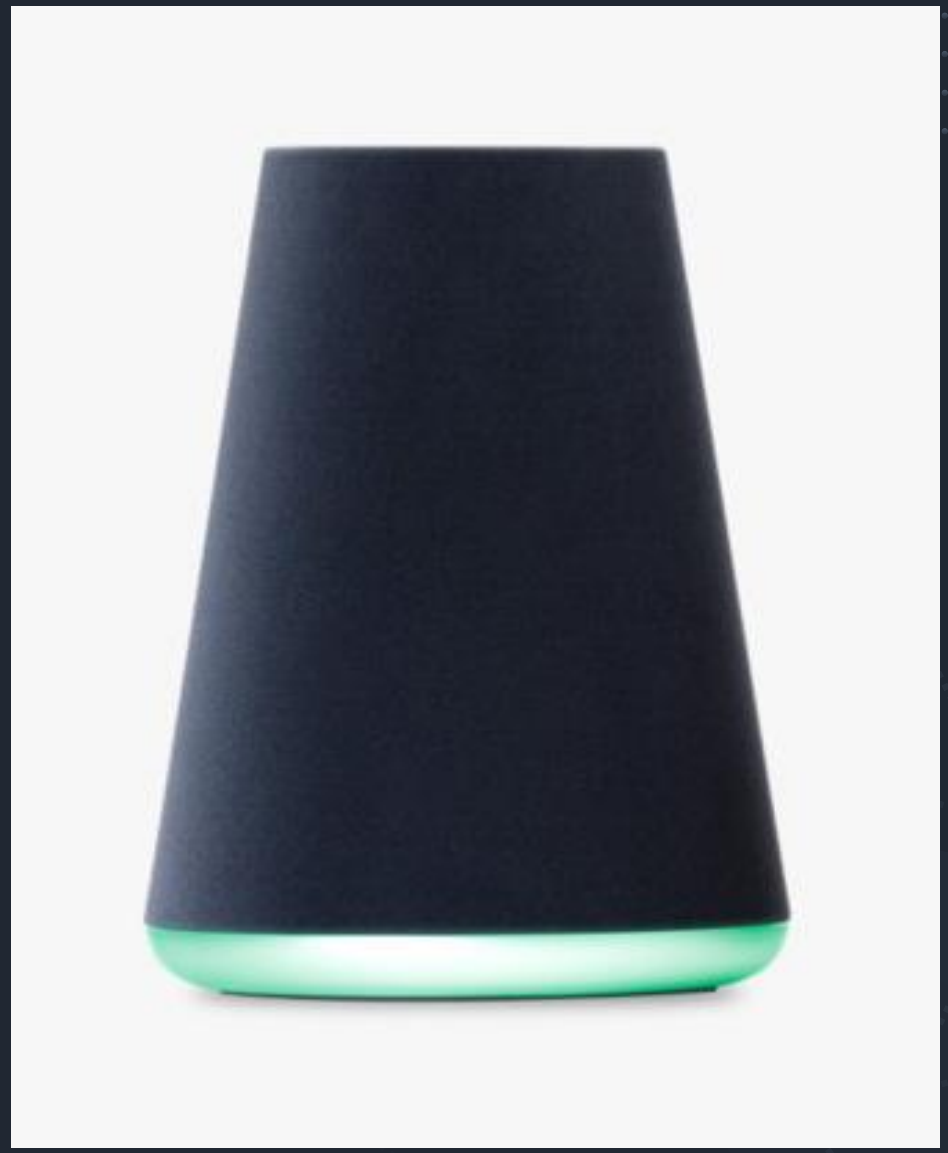
Table 1. Participant demographics and speech rate of screen readers they used. (Stand-alone CA they have experienced: *owned, **tried)

2.2. Method : Apparatus

제품 선택 이유

1) 다른 CA들과는 달리 이 제품은 속도조절 기능지원

2) 날씨 예보, 음악재생, 잡담 등의 기능지원



네이버가 개발한 인공지능 스피커 클로버 웨이브

2.3. Method : Procedure

- 총 20일 간의 평가가 끝난 후, ▲전반적 사용 만족도 ▲실효성 ▲편리성 ▲친숙성 ▲익숙함 등의 지표를 7-point Likert scale 설문조사진행

- 또한 참여자 모두 25~60분 간의 인터뷰 진행

- 인터뷰 내용은 ▲speech rate ▲기존에 본인들이 쓰던 Screen reader와 CA와의 차이점 ▲본 평가를 하면서 본인들이 느꼈던 점 및 개선방안 등으로 이뤄짐

	Day 1 to 10	Day 11 to 20
Group 1 (P1 to P5)	Default-rate CA	Fast-rate CA
Group 2 (P6 to P10)	Fast-rate CA	Default-rate CA

Table 2. Participants' groups and CA usage sequence.

Default-rate : 4.72 SPS (한국인 평균 : 4.82SPS)

Fast – rate : 7.22 SPS

cf) SPS : Syllable Per Sentence

2.4. Method : Data Analysis

앞선 Procedure에서 취합한 인터뷰를 토대로 9개의 main Theme와 29개의 sub Theme로 분류

Overall CA experiences of visually impaired users
1. Enhancing independence 2. Helping visual difficulties 3. Limited functions 4. Limited answers 5. Lack of consideration for people with vision impairments
Positive opinions at the Fast-rate CA
6. Using fast speech rate in their existing screen readers 7. Feeling intimate even at a fast rate. 8. Keeping privacy
Negative opinions at the Fast-rate CA
9. Missing information 10. Distance problem 11. Family members: low vision or sighted people
Perceptions toward CAs at two different speech rates
12. Perceiving the Fast-rate CA as more mechanical thing 13. Perceiving the Default-rate CA as more intimate thing 14. Perceiving the Default-rate CA as a mechanical thing
Expectations toward CAs at two different speech rates
15. Conversations similar to human conversation 16. Conversations at a default-rate CA 17. Accents and nuances
Relationships between user expectations and CA speech rate
18. Efficiency of information acquisition: Fast speech rate 19. Daily conversation: Default speech rate
Screen reader and CA
20. Screen reader as a reading machine 21. CA as a conversation partner
Speech rate control
22. Individual differences for speech rate preference 23. Different speech rate depending on context 24. Different speech rate depending on contents 25. Different speech rate within the same task or contents 26. Fine control 27. Speech rate control for the people with vision impairments
Voice elements in CA
28. Sensitiveness for various voice elements 29. Voice elements

Table 3. Primary themes and subthemes from thematic coding.

3.1. DISCUSSION AND DESIGN IMPLICATIONS : Should Conversational Agents Speak Fast?

▲**결과** : 전반적으로 기본 속도가 더 호응도가 좋았음

▲**결과에 대한 원인 :** 1) 기본적으로 무슨 말 하는지 이해가능

2) 가족의 영향

▲**결과에 대한 INSIGHT :** CA(양방향 소통)는 screen reader(단방향 소통) 과 다르기에, 디자인 설계에 있어서 중요한 요소는 유저와의 관계 형성에 중점을 둬야 함을 제시

(∵ 인터뷰 결과, 대부분의 유저들이 a desire for humanness을 갖고 있음에 근거한 것)

cf) 질문에 대한 답 : *The answer we suggest through the results and discussions of this study is 'the human-like speech rate could be provided as a default for people with vision impairments.'*

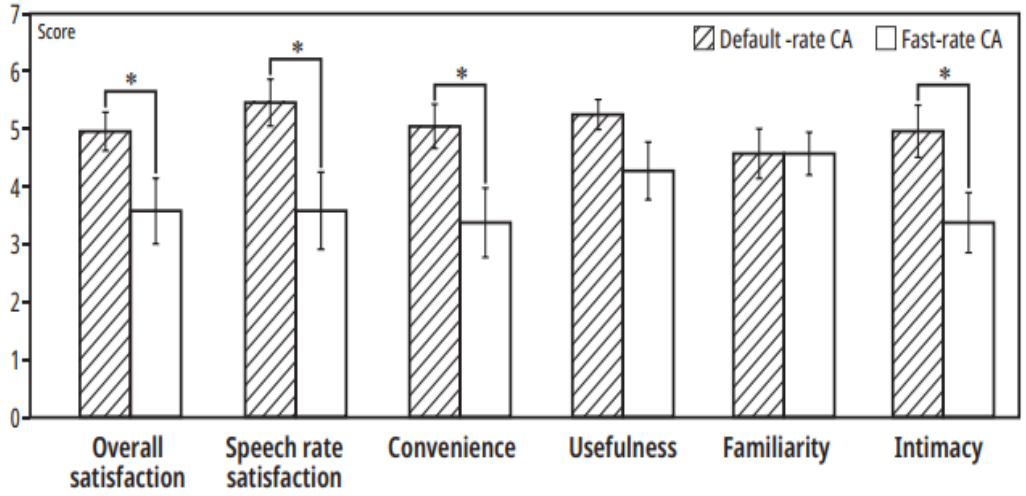


Figure 1. Average survey score for user satisfaction.

	Default-rate CA (Mean + SD)	Fastrate CA (Mean + SD)	z	p
Overall satisfaction	5.00 ± 1.05	3.60 ± 1.78	-1.956	0.050
Speech rate satisfaction	5.50 ± 1.27	3.60 ± 2.12	-2.057	0.040
Covenience	5.10 ± 1.20	3.40 ± 1.90	-2.395	0.017
Usefulness	5.30 ± 0.82	4.3 ± 1.57	-1.496	0.135
Familiarity	4.6 ± 1.35	4.6 ± 1.17	-0.175	0.861
Intimacy	5.00 ± 1.41	3.40 ± 1.65	-2.136	0.033

Table 4. The results of the Wilcoxon signed-rank test for user satisfaction.

3.2. DISCUSSION AND DESIGN IMPLICATIONS

: Speech Rate as a CA Design Element

▲인터뷰 취합 결과 : 1) 사용자들은 상황과 컨텐츠에 따라 다른 속도와 억양으로 말하기를 원함

ex) 상황 : ▲잠에서 깬 직후 ▲외출준비 할 때는 ***빠르게***

컨텐츠 : 뉴스와 책 등의 정보전달의 매체에서 중요함의 여부에 따라 빠르기 다르게

2)▲default 속도에서는 실제 사람 ▲ 빠른 속도에서는 기계와 대화하는 것 같다. 이에 default에 대한 호응

-> 이를 통해 말하기의 속도가 사람의 감정에 영향을 미친다는 것을 도출

3) CA들마다 고유의 특색을 원함

▲ **인터뷰에 대한 INSIGHT :** speech rate는 CA 디자인의 구성요소로서 매우 중요한 것임은 틀림없음. 하지만 이를 보다 user-friendly하게 다가가기 위해서는 특정 문맥 혹은 상황에 따라 ▲tone ▲nuance ▲personality ▲form factors 등도 함께 고려해야함을 제시

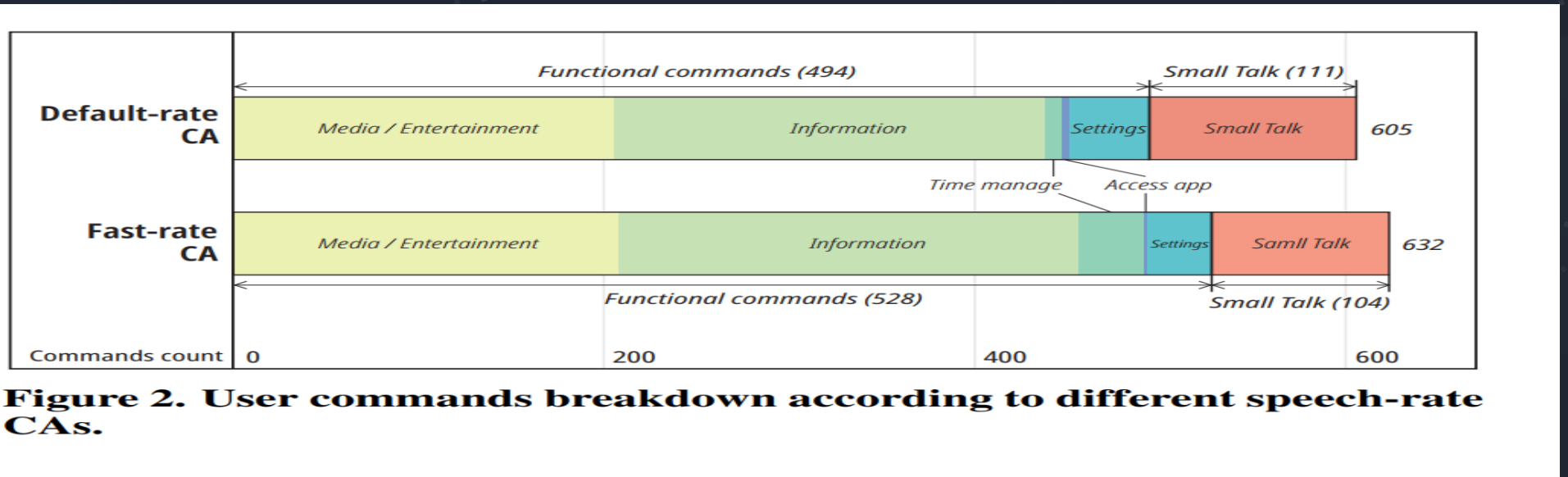


Figure 2. User commands breakdown according to different speech-rate CAs.

4. LIMITATIONS AND FUTURE WORK

LIMITATIONS	FUTURE WORK
시각장애인들만으로 구성	일반인들에게도 확대적용
두 가지 속도(default, fast)로만 측정	-
Total blindness와 low vision에 대한 구분 X	-

THE END