Demonstrating 1D-Touch: NLP-Assisted Coarse Text Selection via a Semi-Direct Gesture

Peiling Jiang*
peiling@ucsd.edu
City University of Hong Kong
Hong Kong, China
University of California San Diego
La Jolla, California, USA

Li Feng
Lfeng256@connect.hkust-gz.edu.cn
City University of Hong Kong
Hong Kong, China
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China

Fuling Sun
fulingsun@ucsd.edu
University of California San Diego
La Jolla, California, USA

Haijun Xia
haijunxia@ucsd.edu
University of California San Diego
La Jolla, California, USA

Can Liu†
canliu@cityu.edu.hk
City University of Hong Kong
Hong Kong, China

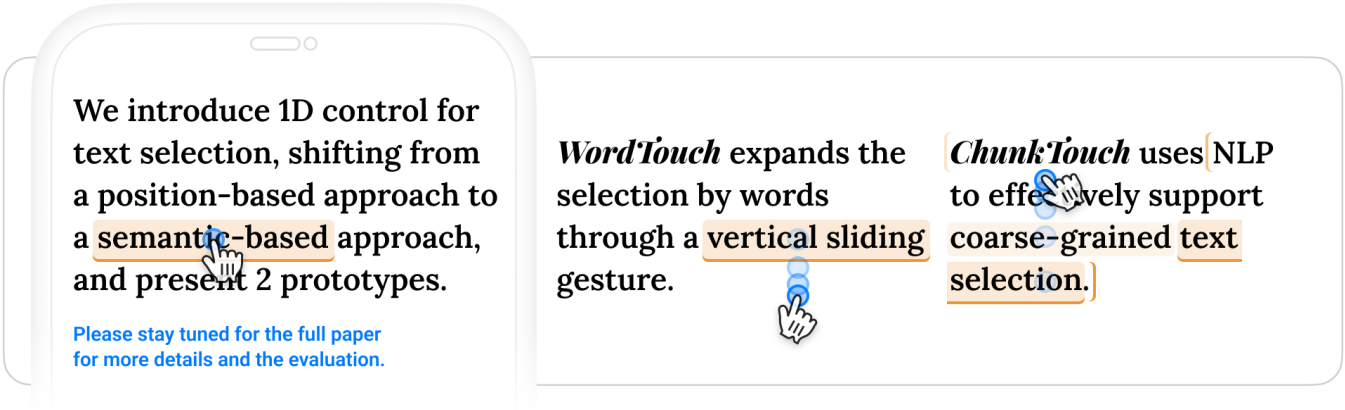

Figure 1: Illustration of two demonstrated 1D-Touch techniques - WordTouch and ChunkTouch.

ABSTRACT

Existing text selection techniques on touchscreen focus on improving the control for moving the carets. Coarse-grained text selection on word- and phrase- levels have not received much support beyond word-snapping and entity recognition. We introduce 1D-Touch, a novel text selection method that complements the carets-based sub-word selection by facilitating the selection of words and larger semantic units. This method employs a simple vertical slide gesture to expand and contract a selection area from a word. The expansion can be by words or by semantic chunks ranging from sub-phrases to sentences, as implemented in two variants of our technique named WordTouch and ChunkTouch. This approach shifts the concept of text selection, away from defining a range by locating the first and last characters, towards a dynamic process of expanding and contracting a textual entity. While the full paper (expected to appear at the ACM ISS 2023) details the evaluation, this demonstration showcases 1D-Touch with a few applications of coarse-grained text selection, to engage the audience in discussions about its effectiveness and applications, as well as its integration with existing character-level selection techniques.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Interaction techniques**.

KEYWORDS

Text selection, Natural Language Processing, Touch interface

ACM Reference Format:
Peiling Jiang, Li Feng, Fuling Sun, Haijun Xia, and Can Liu. 2023. Demonstrating 1D-Touch: NLP-Assisted Coarse Text Selection via a Semi-Direct Gesture. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct), October 29–November 01, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3586182.3615797

*This work was done when the author was an intern at City University of Hong Kong.
†Corresponding author.

UIST '23 Adjunct, October 29–November 01, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0096-5/23/10.
https://doi.org/10.1145/3586182.3615797

1 INTRODUCTION

Text selection through direct manipulation on touchscreens is known to be difficult due to the Fat Finger Problem [7]. The target acquisition task itself consists of defining a range between the start and end positions. The fundamental interface used for this task involves moving a pair of carets to these positions between words or chars. To mitigate the challenge of direct manipulation with low precision input, previous research has focused on introducing *indirection* to text selection by providing indirect control of the cursor or caret positions [1, 2, 4, 5].

Although character-level text selection is important, there are many text manipulation tasks that focus on a larger granularity of text that conveys semantic meanings with words, phrases, clauses, or sentences. Typical examples include highlighting key phrases for note-taking, selecting phrases to be translated, copying/pasting Named Entities (NE) or other semantic segments, etc. Furthermore, there are increasing needs and opportunities for supporting semantic-based text manipulation, with the rapid rise of speech-based multimodal text manipulation, which often operates on the word-level [8], and the breakthroughs in Large Language Models (LLMs) that operates on the tokens (i.e. words).

We present 1D-Touch — a new approach for coarse-grained text selection in the granularity of words and above. Using one-dimensional input, 1D-Touch shifts the concept of text selection, from specifying two 2D positions for defining a range, to selecting one word as the initial target and gradually expanding/contracting it to adjacent words to form a syntactic/semantic object. Users start by selecting the first word with a *direct* touch-hold and then slide up or down to *indirectly* control the expansion of the selection in both directions. Our technique can be set to expand the selection by semantic units in sizes ranging from words, sub-phrases to sentences. We prototyped it by leveraging the constituency parsing [3] in existing NLP tools to chunk text, providing a more gradual and flexible segmentation than basic units that only include words, phrases, sentences, and paragraphs.

To our knowledge, 1D-Touch is the first to introduce *1D control* and *continuous linguistic chunking* to text selection. To help understand the effects of these two novel features, we demonstrate two variants of our technique with the vertical sliding gesture controlling selection expansion by words (WordTouch) and by chunks (ChunkTouch), respectively. While our full paper, expected to be published in Proceedings of the ACM Interactive Surfaces and Spaces (ISS) 2023, details the evaluation and analysis of their performance gain against state-of-the-art, this demonstration focuses on illustrating the interaction and applications of 1D-Touch.

2 1D-TOUCH

This demonstration introduces the use of a one-dimensional sliding gesture for text selection with **1D-Touch**[1]. We implemented two variants of the technique — **WordTouch** and **ChunkTouch**. For both, a long press of 500 ms on a word initiates the initial selection of the word. Without lifting the finger, users could slide either down or up to expand or contract the selection, controlled by the vertical sliding distance. WordTouch expands selection word by word, and ChunkTouch expands in syntactic units derived from NLP constituency parsing. We implemented the prototype as a web application with React.

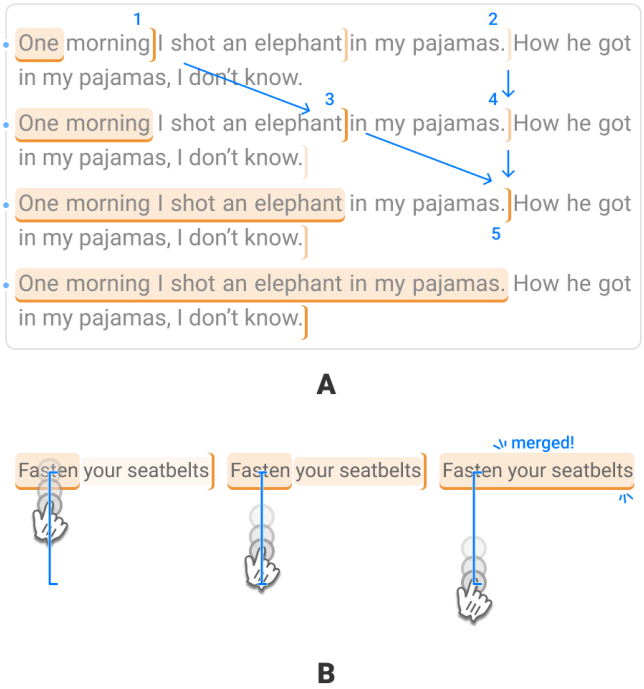

Figure 2: Selection expansion in ChunkTouch and feedback design. (A) The selection expands to the end of syntactic units for each increment. (B) Changing the background provides feedforward of the next expansion unit as the finger slides away from the initial touchpoint.

For ChunkTouch, we used Node.js to run a wrapped version of CoreNLP to perform constituency parsing [3]. Figure 2 illustrates the chunking results, as visualized by the graphical interface with *brackets*. The brackets provide feedback and feedforward for the selection expansion. We consider semantic units of text as chunks of varying sizes that represent meanings at varied levels of abstraction [6]. For this prototype, we choose *word* as the smallest semantic unit and *syntactic unit* as a basic method to chunk text in levels that correspond to semantics.

Furthermore, we implemented supporting features to further improve the usability of the techniques, as introduced below.

Rewinding the overshot selections. Overshoot is a common problem in text selection, where one selects more than they originally planned to select. 1D-Touch techniques support rewinding to fix an overshoot. After overshooting, before lifting the finger, the user can slide backward to rewind the current selection. Rewinding is always by words to allow flexible word-level selection for ChunkTouch.

Clutching with increasing target size. 1D-Touch supports clutching, which allows users to continue adjusting a selection by sliding again. If one shot of sliding did not expand the selection enough to include all target text, users can lift the finger and start another sliding gesture on the *selected* text. The selection becomes increasingly easier as the user clutches, as the target area grows bigger.

[1]An online prototype of the 1D-Touch techniques is deployed at https://public--text-selection.netlify.app/. Please access with a device with a touch screen.

Automatic scrolling for off-screen text. When the selection expands to the bottom or top of the screen, users can continue sliding downwards to grow the selection to the off-screen text. 1D-Touch automatically scrolls the content to keep the end of the selection visible. Auto-scroll for text selection is usually time-based — the user drags a caret to the end of the page, holds it, and waits for the page to scroll to include the content of their interest. However, the waiting time and scrolling distance are hard to predict and control. 1D-Touch, on the other hand, allows the users to directly control the scrolling through sliding distance.

Integration with character-level selection. In addition, we explored ways to integrate our word and chunk selection with character-level selection. This could be achieved by showing carets for character-level adjustment when the user lifts the finger after coarse selections through sliding, as implemented as the default selection technique in existing mobile systems. On the other hand, we demonstrate a new approach that extends our technique by utilizing the horizontal sliding gesture for character-level adjustment.

3 DEMONSTRATION OF APPLICATIONS

To demonstrate 1D-Touch, we implemented a web application to support several use cases that could benefit from coarse-grained text selection. During the demonstration, the audiences use 1D-Touch techniques to perform text selection tasks in the following application scenarios.

Translation. The audience is presented with a piece of text with language jargon. They are able to select phrases or sentences for translation. This scenario highlights the utility of 1D-Touch in quickly selecting semantically meaningful units that usually improve translation results.

Highlighting. The audience is asked to highlight certain text within a lengthy article. They need to select and highlight key arguments related to a particular statement for note-taking.

Speech-based editing. 1D-Touch is well-suited for speech-based text editing scenarios. The audience is able to select text and then speak new content to replace it. Complete semantic phrases and sentences provide more context for speech recognition models compared to individual words, which contributes to better recognition results. We expect ChunkTouch to be utilized to assist text modification when using speech as the input method.

4 CONCLUSION

This demonstration introduces 1D-Touch, a novel approach for text selection that utilizes a one-dimensional semi-direct gesture. The technique improves coarse-grained text selection while complementing the existing carets interface for character-level adjustments. It points towards a promising future for more efficient text interaction by leveraging the fast-developing ability of language AI in understanding text semantics.

ACKNOWLEDGMENTS

This research was supported by the Hong Kong Research Grants Council - ECS scheme under project number CityU 21209419. We thank Shumin Zhai from Google LLC for the insightful discussions.

REFERENCES

[1] Toshiyuki Ando, Toshiya Isomoto, Buntarou Shizuki, and Shin Takahashi. 2018. Press & Tilt: One-Handed Text Selection and Command Execution on Smartphone. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (Melbourne, Australia) *(OzCHI '18)*. Association for Computing Machinery, New York, NY, USA, 401–405. https://doi.org/10.1145/3292147.3292178

[2] Toshiyuki Ando, Toshiya Isomoto, Buntarou Shizuki, and Shin Takahashi. 2019. One-Handed Rapid Text Selection and Command Execution Method for Smartphones. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312850

[3] Gerardo Bort. 2019. CoreNLP for NodeJS. https://github.com/gerardobort/node-corenlp

[4] Vittorio Fuccella, Poika Isokoski, and Benoit Martin. 2013. *Gestures and Widgets: Performance in Text Editing on Multi-Touch Capable Mobile Devices.* Association for Computing Machinery, 2785–2794.

[5] Jianwei Lai, Navid Rajabi, and Elahe Javadi. 2019. A Shortcut for Caret Positioning on Touch-Screen Phones. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) *(MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article 35, 6 pages. https://doi.org/10.1145/3338286.3340146

[6] V.F. Reyna and C.J. Brainerd. 1995. Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences* 7, 1 (1995), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4 Special Issue: Fuzzy-Trace Theory.

[7] Katie A. Siek, Yvonne Rogers, and Kay H. Connelly. 2005. Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In *Human-Computer Interaction - INTERACT 2005*, Maria Francesca Costabile and Fabio Paternò (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 267–280.

[8] Maozheng Zhao, Henry Huang, Zhi Li, Rui Liu, Wenzhe Cui, Kajal Toshniwal, Ananya Goel, Andrew Wang, Xia Zhao, Sina Rashidian, Furqan Baig, Khiem Phi, Shumin Zhai, IV Ramakrishnan, Fusheng Wang, and Xiaojun Bi. 2022. EyeSayCorrect: Eye Gaze and Voice Based Hands-Free Text Correction for Mobile Devices. In *27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 470–482. https://doi.org/10.1145/3490099.3511103