

## Real-time processing of indexical and generic expressions: Insights from, and implications for, COVID-related public health messages

Elsi Kaiser and Jesse Storbeck, University of Southern California <emkaiser@usc.edu>

Much work on real-time referential processing has centered on 3rd person *anaphoric pronouns* (e.g. *she/he*) and has generated foundational insights about antecedent retrieval, salience, and discourse representations. However, the real-time processing of *indexical pronouns*, whose reference changes from one context to another (e.g. *I, you, we*) has received less attention. Indeed, most existing psycholinguistic accounts of pronoun processing are effectively accounts of *anaphoric* pronoun processing, even though indexicals are some of the most frequent and communicatively central parts of language, and semantically fundamentally different from anaphoric pronouns (e.g. Braun 2001). Pronouns with generic reference (e.g. 'you' = people in general, akin to 'one') have also received little attention in real-time processing work.

We use COVID-related health messages to investigate the processing of indexical and generic expressions, with the dual aims of (i) contributing to our understanding of how these experimentally under-researched expressions are processed in real time and (ii) exploring whether the ease of comprehending public health messages related to the COVID pandemic (as measured by reading time) is influenced by type of referring expression.

Prior work on 1st and 2nd person pronouns in COVID messages (and health messages generally) is limited, with mixed results. E.g. Tu et al. (2021) found that COVID stay-at-home messages with *you* are more effective than ones with *we* when it comes to shaping people's self-reported likelihood of staying at home vs. going to a friend's party in a hypothetical scenario. Kaiser (2021)'s work on *you, we* and *people/everyone* in subject position found that COVID messages about masks and social distancing with *people/everyone* were rated more convincing by democrats, while non-democrats showed no clear pronoun-type effects.

However, these studies did not measure processing ease. Given the value of easily-understood health messages (e.g. CDC's *Health Communication Playbook*), identifying differences in processing ease of different expressions has theoretical and applied relevance.

**Experiment.** Using self-paced reading, we tested COVID-related behavior recommendations (ex.1). The study had 39 targets and 48 native English-speaking, U.S.-based participants recruited via Prolific who participated via PClbex (Zehr & Schwarz 2018). We manipulated referring expression type (*we, you, people*; within-subjects Latin Square design). The referring expression was preceded by a short preamble phrase (eg. 'on account of the pandemic') and followed by an auxiliary verb, the main verb and the rest of the sentence.

(1) On account of the pandemic, we/you/people should get the vaccine to prevent further spread of COVID-19.....

COVID messages with deontic modality like (1) have the advantage of allowing for minimal triplets where the communicative function of *we/you/people* is as similar as possible: in messages like (1), the communicative goal is constant regardless of which of the forms is used. (This is not the case in examples like 'We go to Italy in the summer' vs. 'People go to Italy in the summer.' Thus, health messages are well-suited for comparing the different expressions.)

In (1), presented out of context, *you* is (in principle) ambiguous: On an indexical interpretation, it refers to the addressee (e.g. Brunyé et al. 2009 on *you* triggering a participant perspective). On a generic interpretation, *you* refers to people in general (like 'one'). *We* is also potentially ambiguous between indexical and generic readings (Holmberg 2017), but its generic use is less frequent than generic *you*. In contrast, *people* is not indexical and only receives a generic-type interpretation. We test 3 hypotheses about the processing ease of these forms:

**Indexicality hypothesis:** Indexically-interpreted pronouns refer to highly salient referents, and do not require evoking/constructing a new discourse referent or even a generic

operator/referent. If this special property of indexicals is hard-wired into the representation of these forms, then we might expect any expressions that *can* in principle have an indexical interpretation to be easier to process than a form that *can never* be indexical: you, we < people. The special status of indexicals receives initial indirect support from the results of Warren & Gibson (2002, 2005) on the processing of indexicals in a different context.

**Perspective-taking hypothesis:** Although indexicals refer to salient referents, they are also perspective-sensitive: The referent of indexical *we* or *you* depends on who the speaker and the addressee are. If this perspective-sensitivity is hard-wired into our processing of certain pronouns (regardless of context), then – in light of prior work suggesting that perspectival processing is cognitively costly (e.g. Keysar et al. 2000, Ferguson et al. 2017) – *you* and *we* may be harder to process than *people* which has no indexical component (people < you, we).

**Genericity hypothesis:** Given that all 3 forms can receive generic interpretations in contexts like (1), this could render the indexical readings of *we* and *you* irrelevant/unavailable in this context. If so, we may see no differences between the three forms (you = we = people).

**Results:** Reading time (RT) data is in Fig.1. Mixed-effect regression models were used to analyze log-transformed RTs. Overall, messages with *people* elicit RT slowdowns relative to messages with *you* and *we*, which do not differ – supporting the **Indexicality Hypothesis**.

Specifically, there are no effects of referential form before the critical region, and no effects at the critical region itself (*you/we/people*, “0” in Fig.1). At spillover region 1, *people* conditions are marginally slower than the *you* ( $p=0.061$ ) and *we* ( $p=0.0768$ ) conditions. At spillover region 2, there are no significant differences. At spillover region 3, *people* conditions are significantly slower than *you* ( $p<.005$ ) and *we* ( $p<.005$ ) conditions. At spillover region 4, there is still a marginal slowdown in *people* conditions relative to *you* conditions ( $p=0.09$ ) but no other differences. Spillover region 5 shows no significant differences.

(As the word *people* is longer and less frequent than *we* or *you*, the marginal slowdowns in spillover region 1 may be due to these surface factors. Crucially, these effects are not significant in region 2. Thus, it seems reasonable to view the slowdown in region 3 as a meaningful indication that *people* sentences trigger slowdowns relative to *you* and *we* for reasons independent of word length/frequency.)

We have also conducted a

**between-subjects version** of this study ( $n=48$  new people) with the same items and method but with referring expression type manipulated between-subjects, which **replicates** the finding that *people* conditions are read more slowly than *you* or *we* conditions.

**Conclusions.** This study takes initial steps to explore the real-time processing of non-anaphoric pronouns by focusing on indexical and generic forms in COVID health messages. To the best of our knowledge, this is the first real-time study to test how -- in COVID health messages -- different forms (*you*, *we*, *people*) impact reading time, which we take to reflect ease of processing. Our results point to an increased processing load in messages with the non-indexical form *people* (relative to pronouns *we* and *you*) which (i) we interpret as providing initial support for the Indexicality Hypothesis, and which (ii) also has practical implications for the construction of easily-understood public health messages.

