

An Empirical Study on Vague Deictic Temporal Adverbials

Svenja Kenneweg¹, Brendan Balcerak Jackson¹,
Jörg Deigmoeller², Julian Eggert², Philipp Cimiano¹

¹Bielefeld University, Germany

{skenneweg, bbalcerakjackson, cimiano}@techfak.uni-bielefeld.de

²Honda Research Institute Europe, Germany

{Joerg.Deigmoeller, Julian.Eggert}@honda-ri.de

Abstract

Temporal adverbial phrases such as *recently* and *some time ago* have a special function in communication and temporal cognition. These adverbials are deictic, in that their meaning is tied to their time of utterance; and they are vague, in that the time periods to which they apply are under-specified in comparison to expressions such as *yesterday*, which precisely indicates the day before the day of utterance. Despite their vagueness, conversational participants have a mental image of when events described using these adverbials take place. Our study aims to quantify this mental model in terms of fuzzy or graded membership. To achieve this, we investigated the four English temporal adverbials *recently*, *just*, *some time ago* and *long time ago* as applied to types of events with different durations and frequencies, by conducting surveys to measure how speakers judge the different adverbials to apply in different time ranges. Our results suggest that it is possible to represent the meanings of deictic vague temporal adverbials geometrically in terms of graded membership within a temporal conceptual space.

Keywords: Temporal adverbials, Meaning representation, Deixis, Vagueness, Cognitive semantics

1. Introduction

Temporal adverbial phrases such as *recently*, *soon*, *some time ago* and *in a while* play a distinctive role in communication and cognition pertaining to time. For example,

(1) I just had my birthday.

If person A utters (1) to person B, B is likely to assume that A's birthday was in the past several days, and might respond by congratulating A. But suppose A utters the following instead:

(2) My birthday was some time ago.

B is unlikely to congratulate A, and B would be surprised to learn that A's birthday was in fact only two days prior.

Temporal adverbials like these are instances of *deixis*, or context-sensitivity: the range of times to which the adverbial applies depends on the time at which it is uttered, and different utterances apply to different ranges of times. Other deictic expressions are pronouns such as *I* and *you*, demonstratives such as *this* and *that*, and locative expressions such as *here* and *there*.

These adverbials are also instances of *vagueness*. For a given utterance of (1), for example, there is a range of times prior to the time of utterance which would not clearly be correctly described as times when the speaker had 'just' had her birthday, but which would also not clearly be described as time when she had *not* 'just' had her birthday. Other ex-

amples of vague expressions are gradable adjectives such as *tall* and *old*, adverbs such as *quickly* and *loudly*, and nouns such as *pile* and *heap*.

Not all vague temporal adverbials are deictic, e.g. *just before Christmas 2023*. And not all deictic temporal adverbials are vague, e.g. *exactly 37 minutes ago*. Our focus here is on temporal adverbials that are both deictic and vague. We take a cognitive approach, seeking to understand how the mind conceptualizes times and events in terms of vague deictic temporal adverbials.

We present an initial empirical study conducted with 100 adult participants concerning the past temporal adverbials *just*, *recently*, *some time ago*, and *long time ago*. For each adverbial, subjects were given a range of scenarios involving past events and were asked to evaluate the extent to which the expression applies in the scenarios. The aim of the study was to measure how language users represent events, both in terms of the time of utterance (deictic aspect) and in terms of fuzzy membership (vague aspect). In future work this study can lay the foundation in developing a computational model of vague temporal adverbials.

2. Related Work

2.1. Vagueness

Vagueness is standardly characterized by the existence of borderline cases: an expression is vague just in case there are cases in which ordinary speakers judge that the expression neither clearly applies nor clearly fails to apply. The

adjective ‘tall’, as applied to persons, counts as vague because there are borderline cases of tall people, i.e. people who neither clearly count as tall nor clearly count as not tall. Numerous logico-linguistic frameworks have been proposed for making sense of borderline cases and vagueness; useful surveys are provided in [Keefe and Smith \(1997\)](#); [Keefe \(2000\)](#); [Kennedy \(2019\)](#); [Solt \(2015\)](#).

The most directly relevant framework for the present study is the *fuzzy* approach to vagueness ([Damerou \(1977\)](#); [Hájek \(1998\)](#); [Zadeh \(1965, 1973\)](#)). While classical semantics treats membership in a class such as *tall people* as an all-or-nothing matter, fuzzy approaches distinguish different degrees of membership on the closed interval $[0,1]$. A borderline case of a tall person is someone whose degree of membership measures somewhere in the middle of this interval. Early fuzzy approaches tended to interpret graded membership extensionally as a relationship between the expression and entities in extra-linguistic reality (e.g. actual and possible people of different heights). Here we construe graded membership in cognitive terms, as representing the way language users mentally represent reality (see [Hersh and Caramazza \(1976\)](#); [Douven et al. \(2013\)](#)); we discuss this further in Section 2.2.

The vast majority of literature on vagueness either abstracts from specific categories of expressions, or else focuses on vagueness in the adjectival domain. Very little work focuses specifically on vagueness in temporal adverbials, and virtually none we could find investigates vague temporal adverbials experimentally. A notable exception is [Van Jaarsveld and Schreuder \(1985\)](#), an empirical study of temporal adverbials in Dutch. Their findings suggest that the range of times for which a speaker is disposed to apply a temporal adverbial to an event is influenced by the subject’s beliefs about the frequency and duration of events of that type. This study and its methodology provided the starting point for the present work.

2.2. Conceptual Space Semantics

Here we adopt a cognitive perspective on semantics that takes linguistic expressions to correspond to concepts or ways of mentally representing reality. Particularly relevant for the present study is the *conceptual spaces* framework of [Gärdenfors \(2014\)](#), in which concepts are represented geometrically, as regions in spaces that are defined in terms of one or more representational dimensions. For example, humans represent color using a three-dimensional conceptual space defined in terms of the dimensions *hue*, *saturation*, and *brightness*. Color concepts, and the meanings of color terms, correspond to regions of this con-

ceptual space that have certain formally specified properties ([Gärdenfors \(2014\)](#)); the location of a given object within a region indicates the way the subject represents its color. Other concepts correspond to regions in conceptual spaces constructed from dimensions corresponding to other qualities, such as spatial or temporal extent, auditory experience or taste, and different kinds of motion and action.

As [Douven et al. \(2013\)](#) and [Decock and Douven \(2014\)](#) argue, conceptual spaces provide a natural way of interpreting the graded membership relation employed by the fuzzy approach to vagueness. The metric of a conceptual space makes it possible to identify *prototypes* for a concept, subregions that have distinguished positions within the region corresponding to the concept. The degree to which an object belongs to the concept can then be calculated using its distance from the prototypes for the concept together with its distance from the prototypes of adjacent concepts in the space. A borderline case of *red*, for example, might have a degree of membership of 0.4 because it is just about as far away from the prototypes for red as it is from the prototypes for orange. The mathematical details can become somewhat complex, especially for high-dimensional conceptual spaces (see [Decock and Douven \(2014\)](#) for discussion) but they are not necessary for present purposes.

3. Experimental Setup

We devised a set of online surveys to measure language users’ representations of events in terms of a graded or fuzzy relation to the time of utterance. Our study centered on the four temporal adverbials *recently*, *just*, *long time ago* and *some time ago* as applied to hypothetical events involving an imaginary person named Tom. In light of the results in [Van Jaarsveld and Schreuder \(1985\)](#), which suggest that the applicability of a temporal adverbial is influenced by the duration and frequency of the associated event, we designed five separate surveys for five types of events with different durations and frequencies: Birthday, Brushing Teeth, Marriage, Vacation and Year Abroad. Brushing Teeth has the shortest duration, which is typically only a few minutes. In contrast, Year Abroad has a significantly longer duration of one year. The durations of the other events fall somewhere in between these extremes. The frequency of events follows a comparable reversed hierarchy: Brushing Teeth happens daily and has therefore the highest frequency, while Year Abroad happens only a few times, if at all, in one’s life. The frequencies of the other events lie in between. Each survey queried subjects about a series of English sentences containing a temporal adverbial

	<i>Just Recently</i>	<i>Some time ago Long time ago</i>
Brushing Teeth	5 min.	1 hour

	1 day	3 days
Birthday	1 day	1 week

	3 months	11 months and 3 weeks
Vacation	1 day	4 days

	4 months	1 year
Marriage Year Abroad	4 days	1 month

	1 year	8 years

Table 1: The nearest and furthest $\langle timeSpan \rangle$ point surveyed for each event type and each temporal adverbial.

applied to a past event (see section 3.1). Participants were asked to rate the applicability of the temporal adverbial in each case on a 5-point Likert scale; "Doesn't apply"; "Barely applies"; "Partially applies"; "Mostly applies"; "Completely applies". Each survey was administered to 100 participants via the Prolific platform, and each participant received a small compensation for completing the survey.¹ All participants were adults with English as their native and primary language.

3.1. Survey sentences

Each of the surveys had one survey page for each of the four temporal adverbials, with 7 test sentences on each survey page. Test sentences were constructed using an event description and a $\langle timeSpan \rangle$ indicating the time passed since the occurrence of the event, followed by a statement applying the temporal adverbial to the event. For example, for the event type Birthday, the $\langle timeSpan \rangle > 1$ day and temporal adverbial *recently*, the test sentence is: "Tom's Birthday was 1 day ago. Statement: Tom had his Birthday recently."

The range of values for $\langle timeSpan \rangle$ was chosen separately for each type of event based on our own experience. For each event type we maintained one range of $\langle timeSpan \rangle$ for *just* and *recently*, which concern times closer to the present, and a different range for *some time ago* and *long time ago*, which concern more distant time points. The range for $\langle timeSpan \rangle$ for each temporal adverbial and event is shown in Table 1.

¹<https://www.prolific.com/>

4. Results

Figure 1 shows the median and interquartile ranges (highlighted in brighter colors) of all survey data for each of the temporal adverbials studied, with all five event types plotted together for comparison. The Likert scale points from "Doesn't apply" to "Completely applies" were numerically encoded from 1 to 5. Subsequently we applied a linear transformation to map the values onto a range of 0 to 1, representing degree of membership.

We first compare the plots for the different temporal adverbials, starting with a comparison of *just* and *recently* (Figures 1(a)-(b)). Membership values for both adverbials are highest when the event has just occurred and decrease as the distance into the past increases. However, the decrease is more pronounced for *just* than *recently* across all event types. For example, although the membership value for a marriage that took place four days ago is 1 for both adverbials, speakers judge *recently* to apply to a marriage that took place 4 months ago more strongly than *just*. Turning to *some time ago* and *long time ago* (Figures 1(c)-(d)), we see that the initial values for *some time ago* are higher than those for *long time ago*. For Brushing teeth (which is hidden by the Time Axis) and *some time ago* the values rise to a peak before falling off again as we move further into the past. Our data does not show such a peak for the other events but it is possible that it lies outside the range of times we surveyed. An indication of this is that the interquartile range of Birthday and Vacation is higher at the end for *some time ago*. This suggests that *some time ago* corresponds to a segment of the past with vaguely defined start and end points, while *long time ago* corresponds to any time further in the past than a vaguely defined starting point.

When we compare the different event types across all temporal adverbials, we can see that the arrangement of the event types is always the same: first Brushing Teeth, then Birthday, Vacation, and Year Abroad, and then finally Marriage. The rate of increase or decrease in membership values is always slowest for Marriage and fastest for Brushing Teeth.

5. Discussion

Our results have a natural interpretation within the conceptual spaces framework. [Sinha and Gärdenfors \(2014\)](#) hypothesize a conceptual space defined in terms of *D-time* (deictic time), which orders points in terms of their distance (forwards or backwards) from the present. The meanings of deictic adverbials can be represented as regions in this space, and our results can be understood as helping to map the contours of those regions. This also provides a natural framework for repre-

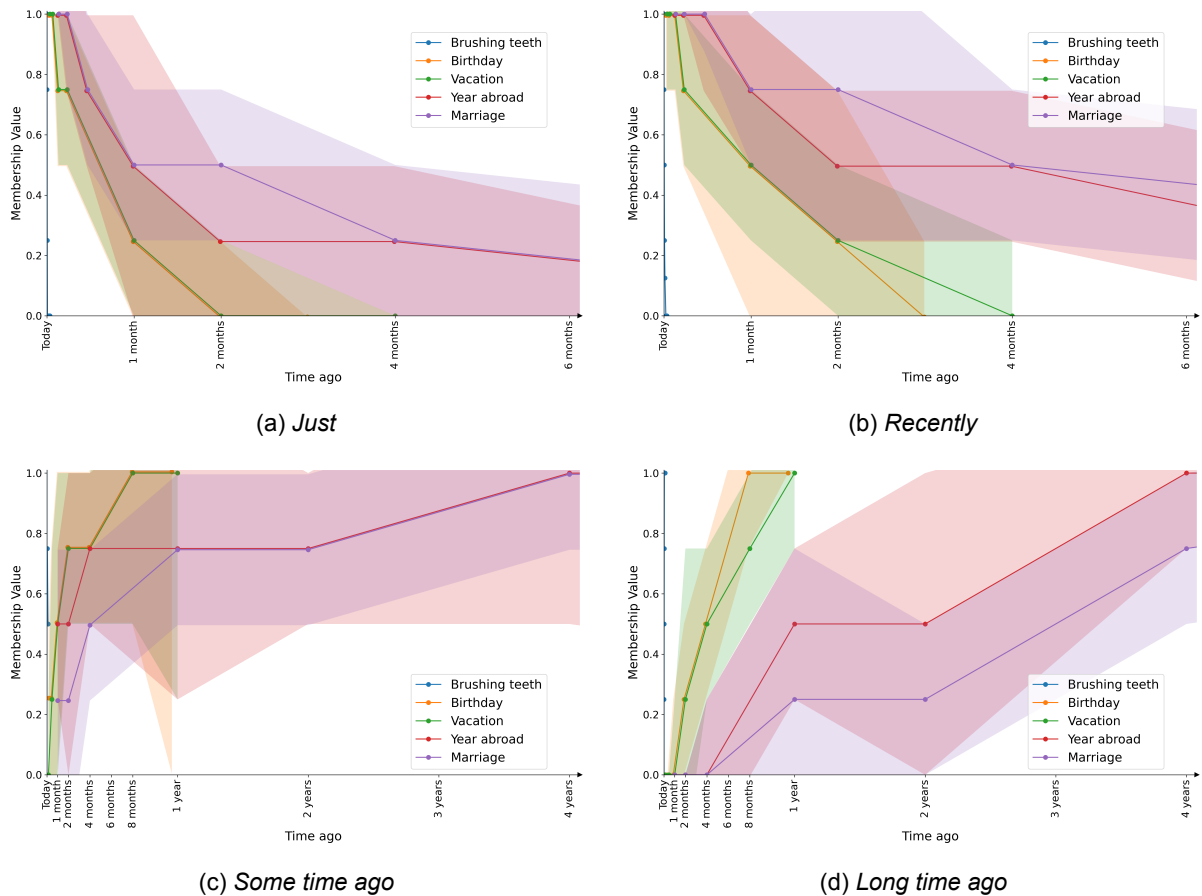


Figure 1: Median survey data and their interquartile ranges (shown in brighter colors) for the 4 evaluated temporal adverbials and 5 events.

senting vagueness in temporal adverbials in terms of graded membership, as sketched in Section 2.2. Gärdenfors (2014) conjectures that the meanings of (nearly) all simple natural language expressions correspond to convex regions of conceptual space. Within a graded membership framework, convexity requires that all the points on the path between any two points that belong to the region to at least membership value α also belong to the region to at least value α . All the regions of D-time measured are convex in this sense.

Our findings are also consistent with the earlier study of Van Jaarsveld and Schreuder (1985) that the interpretation of a temporal adverbial is influenced by the duration and frequency of the event. For each adverbial, all five event types have the same curve progression. (The time scale in Figure 1 makes this difficult to see for Brushing Teeth.) This suggests that the primary effect of combining an adverbial with different event types is to stretch or compress the corresponding region of conceptual space, rather than to produce significant changes in its shape.

Moreover, the methodology used here establishes natural hierarchies among times according to whether a given temporal adverbial, combined

with a given event type, applies more clearly or definitely at one time or another. In principle, this can put us in a position to determine the most probable time of occurrence of an event given its linguistic representation in terms of temporal adverbials; meaning representations that support this kind of inference would be especially valuable for computational implementations.

6. Conclusion & Future Work

We presented a study that measured language users' representations of events in terms of a graded relation to the time of utterance, in order to simultaneously capture the deictic and vague character of the temporal adverbials selected. Our meaningful results provide confirmation that the survey method we employed is suitable for this purpose. Our results are consistent with the work of Van Jaarsveld and Schreuder (1985), and they have a natural interpretation within the conceptual spaces semantic framework.

Our eventual aim is to develop a computational model for applications such as human-robot interaction. Given an event type and a temporal adverbial, such a system should be able to recover

a time span in which the event is most clearly represented as occurring; it should also be able to choose appropriate temporal adverbials for communicating temporal information. Reaching this point requires much more work. A first step is determining exactly how event duration and frequency affect temporal adverbial distribution over deictic time, an additional step being to explore other influencing factors. Other possible influencing factors can be identified by comparing events with the same durations and frequencies, to check for differences in mental representations of the corresponding time spans that cannot be entirely accounted for in terms of the influence of duration and frequency.

Given the recent progress in large language models, another important next step is to check whether they interpret vague deictic temporal adverbials in a way that is consistent with human understanding. This will help determine what role large language models can play in computational modeling of temporal communication and cognition.

This paper focused on deictic temporal adverbials and events. Another area of focus for future work could be on relative adverbials that relate one event to another, such as *Tom was on vacation just before his birthday*.

7. Optional Supplementary Materials

7.1. Limitations

The survey was limited by the number of questions that participants would be likely to answer seriously. According to [Bowling et al. \(2021\)](#), the proportion of unfocused random answers in online studies is less than 1% for studies with no more than 33 questions. Per survey, there are 28 questions about the vague temporal adverbials, two demographic questions and one question about the survey difficulty. Due to budget constraints, there were limits to the number of survey participants and the number of surveys that could be carried out.

7.2. Ethics

An ethics proposal for the surveys was submitted to the ethics committee of the Bielefeld University and was found to be ethically unobjectionable. This complies with the ethical guidelines of the German Psychological Association and the Professional Association of German Psychologists.

8. Bibliographical References

- Nathan A. Bowling, Anthony M. Gibson, Joseph W. Houpt, and Cheyna K. Brower. 2021. *Will the questions ever end? person-level increases in careless responding during questionnaire completion*. *Organizational Research Methods*, 24(4):718–738.
- Fred J Damerau. 1977. On “fuzzy” adjectives.
- Lieven Decock and Igor Douven. 2014. What is graded membership? *Noûs*, 48(4):653–682.
- Igor Douven, Lieven Decock, Richard Dietz, and Paul Égré. 2013. Vagueness: A conceptual spaces approach. *Journal of Philosophical Logic*, 42:137–160.
- Peter Gärdenfors. 2014. *The geometry of meaning: Semantics based on conceptual spaces*. MIT press.
- Petr Hájek. 1998. *Metamathematics of fuzzy logic*. Springer.
- James A Hampton. 2007. Typicality, graded membership, and vagueness. *Cognitive Science*, 31(3):355–384.
- Harry M Hersh and Alfonso Caramazza. 1976. A fuzzy set approach to modifiers and vagueness in natural language. *Journal of Experimental Psychology: General*, 105(3):254.
- Rosanna Keefe. 2000. *Theories of vagueness*. Cambridge University Press.
- Rosanna Keefe and Peter Smith. 1997. Introduction: theories of vagueness. In Rosanna Keefe and Peter Smith, editors, *Vagueness: a reader*. MIT press.
- Christopher Kennedy. 2019. Ambiguity and vagueness: An overview. In Claudia Maienborn, Klaus von Heusinger, and Paul Portner, editors, *Semantics: lexical structures and adjectives*. Walter de Gruyter.
- Charles I. Mosier. 1941. *A psychometric study of meaning*. *The Journal of Social Psychology*, 13(1):123–140.
- Terence Parsons. 1990. Events in the semantics of english: A study in subatomic semantics.
- Chris Sinha and Peter Gärdenfors. 2014. Time, space, and events in language and cognition: a comparative view. *Annals of the New York Academy of Sciences*, 1326(1):72–81.
- Stephanie Solt. 2015. Vagueness and imprecision: Empirical foundations. *Annu. Rev. Linguist.*, 1(1):107–127.
- H.J. Van Jaarsveld and R. Schreuder. 1985. Implicit quantification of temporal adverbials. *Journal of Semantics*, 4:327–339.

Frank Vlach. 1993. Temporal adverbials, tenses, and the perfect. *Linguistics and Philosophy*, 16(3):231–283.

Lotfi Zadeh. 1965. Fuzzy sets. *Information and Control*, 8:338–353.

Lotfi Zadeh. 1973. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on systems, Man, and Cybernetics*, (1):28–44.