# Natural Language Generation and Fuzzy Sets: An Exploratory Study on Geographical Referring Expression Generation

Alejandro Ramos-Soto *, Nava Tintarev†, Rodrigo de Oliveira‡, Ehud Reiter‡, Kees van Deemter‡

* Centro Singular de Investigación
en Tecnoloxías da Información (CiTIUS),
Universidade de Santiago de Compostela
alejandro.ramos@usc.es

† Department of Computing
and Informatics,
University of Bournemouth
ntintarev@bournemouth.ac.uk

‡ Department of Computing Science,
University of Aberdeen
rodrigodeoliveira@abdn.ac.uk
e.reiter@abdn.ac.uk
k.vdeemter@abdn.ac.uk

*Abstract*—We explore how the problem of uncertainty and imprecision in natural language generation (NLG) could be addressed through the use of fuzzy sets. We propose bringing together standard empirical procedures for knowledge acquisition in NLG and computing with words/perceptions related techniques (with a special focus on linguistic description of data) to address an open challenge in NLG: the generation of geographical referring expressions. Following this methodology, we present an exploratory experiment which provides some insights about how human subjects refer to geographical expressions and discuss how the obtained results might relate to the use of fuzzy sets.

## I. INTRODUCTION

Research fields where the management of uncertainty and imprecision play an important role are a potential target for usage of disciplines which have traditionally addressed this problem, such as probability theory or fuzzy logic. One such field is natural language generation (NLG), which deals with the problem of how texts in human language can be created by computers automatically. Texts generated by NLG systems typically contain words and expressions in natural language whose semantics are defined in a crisp way. However, words and concepts used in written language and speech often have an unclear or vague meaning. This motivates considering alternative approaches for NLG which are able to manage the imprecision or vagueness of the terms and concepts used as part of the language generation process.

In this context, the problem of uncertainty in NLG is an open challenge which has not been fully addressed. Examples of NLG work in uncertainty include the proposals by van Deemter [1], where the use of referring expressions involving gradable properties and the practical implications of vague expressions in NLG are explored [2]; an approach by Power and Williams which deals with numerical approximations to describe proportions at different levels of precision [3]; and the extension of the data interpretation and microplanning stages in the BABYTALK family of systems to deal with uncertainty in temporal relations, by Portet and Gatt in [4].

Since uncertainty and vagueness in NLG is a topic of interest for researchers in the field, it does not come as a surprise that NLG has gained attention in recent times from the fuzzy community working on linguistic summarization

(also known as linguistic description) of data. This concept is distinct from the specific sub-field of automatic summarization in computational linguistics, which focuses on creating a shorter text outlining key points from a longer, or several, pieces of text. To avoid confusion, we will refer to "linguistic description of data" instead.

To the best of our knowledge, the first reference from this field to NLG was made by Kacprzyk et al. in [5], while [6] explored for the first time the feasibility of a relationship between linguistic description of data and NLG. Since then, many subsequent research papers in this field have referenced NLG papers as examples of real applied systems (e.g. [7]–[10]).

In fact, a well established opinion in this field is that, in order for linguistic description of data to be successful as a discipline which can be applied to real problems, it should be used in conjuction with NLG techniques. Recent work has further emphasized this idea in more depth [8], [9], while the weather forecast generation system GALiWeather [11], which includes both fuzzy and NLG components, is a real practical example of this trend. Following the same trend, but in opposite direction, Gatt and Portet present from an NLG perspective a model based on fuzzy temporal constraint networks and experimental data from three languages to address the problem of generating uncertain temporal expressions [12].

In this regard, our aim is to explore from a practical point of view how fuzzy set techniques can be applied to model and solve specific challenges on the NLG side, hopefully improving the results obtained with traditional approaches. This paper explores the feasibility and implications of approaching through the use of fuzzy sets the problem of generating geographical referring expressions (GREs) for NLG systems that make use of geographical data.

## II. GEOGRAPHICAL REFERRING EXPRESSION GENERATION

The purpose of an NLG system is the conversion of some input data into an output text. However, as in most computational processes, this task can be decomposed into a number of substasks that address specific problems within the

Fig. 1. UK map with the relevant region highlighted in yellow.

global task. One such depiction of this task division is the well-known and widely accepted architecture proposed by Reiter and Dale in [13], which describes an NLG system as a pipeline composed of several consecutive tasks or modules dealing with different respects of the natural language generation problem. Among them, we will focus in what follows on the problem of referring expression generation (REG), a topic actively researched within the NLG community.

A referring expression (RE), in linguistics, is any noun phrase, or surrogate for a noun phrase, whose function in discourse is to identify some individual object (thing, being, event...) [14]. To generate referring expressions in simple cases may seem trivial (e.g. two chairs which can be distinguished according to their color: "the white chair" or "the black chair"). However, given the wide spectrum of application domains for NLG systems, it is rather common that generating REs requires quite sophisticated approaches [15]. That is the case for NLG systems producing texts from geographical data [16]–[18], where geographical referring expressions (GREs) need to be generated.

One of the challenges of describing geo-tagged events in natural language is linking non-spatial attributes of events to their underlying geography [16]. Data sets that represent events ocurring over a geography have traditionally included latitude-longitude values as the only geographical information, so when automatically generating references to the location of events, a means must be provided to translate coordinate values into more human-friendly, natural language expressions such as "north" or "coast".

For example, assume an NLG system wishes to communicate, in a weather forecast, that there is a chance of rain in the rectangular region between (lat 57.743793°, lon -2.596026°) and (lat 56.719289°, lon -1.744081°). Could this region (see Fig. 1) be described as *northeastern Great Britain*, *northeastern UK*, *northeastern Scotland* or *east Aberdeenshire*?

Turner et al. employed for the NLG system RoadSafe established techniques from the Geographic Information Systems (GIS) community, in which they partition the underlying geography of events using several different schemes, or spatial *reference frames* (or *layers* in GIS terms), which

in turn are composed of non-overlapping partitions (which Turner and colleagues refer to as *descriptors*). For instance, the reference frame Direction is composed of the descriptors "northeast", "southwest", etc., and the frame Coastal Proximity is composed of descriptors "coast" and "inland". Once descriptor boundaries are numerically defined in each frame using latitude-longitude values, each data point can now be linked to a set of descriptors (e.g. "southwest" and "coast") and the job of the referring expression generator is to select a set of descriptors that describe the location of the subset of points that represent the event.

The major issue with this approach is that it assumes that boundaries between descriptors are crisp, so even if one data point is only slightly below another point, one could be computed as "north" and the other one as "south". This very strict approach to partitioning the space seems to be different from what humans naturally do: we seem to be more flexible with boundaries, if we create boundaries at all. This issue is a strong incentive for looking at other approaches which allow to model and manage imprecision or uncertainty in natural language terms, such as fuzzy sets.

### III. USING FUZZY SETS FOR GENERATING GREs

Although the usage of fuzzy set techniques (especially those related to computing with words [19] and linguistic description of data, including linguistic variables, combination operators or fuzzy quantifiers) in the problem of generating GREs seems rather sensible and intuitive, we strongly believe its viability should be assessed and studied empirically first. In this regard, we propose the following methodology (see Fig. 2):

1) Study specifically how georeference is applied by humans, i.e., how people refer to different geographical reference frames.
2) Obtain empirical fuzzy definitions of the primitive geographical descriptors (e.g. "north", "coast", "inland", etc.).
3) Study how the primitive geographical descriptors can be combined in order to obtain more complex descriptors (e.g. "south inland", "northeast coast"...). If fuzzy quantified sentences [20], [21] were to be used either explicitly (e.g. "in some places west inland", "in most of the southern coast") or implicitly (as a way for making flexible the fulfillment of some criteria), obtain empirical definitions for quantifiers and study which quantification models perform better experimentally.
4) Implement the referring expression generation algorithm.
5) Evaluate the fuzzy REG algorithm against a standard approach. Does the fuzzy algorithm produce more appropriate geographical descriptions than previous crisp approaches, according to human subjects?

With 1), we intend to study how human subjects use geographical descriptors, which features they take into account and if results justify the need to model them through fuzzy sets. Step 2) consists in performing experiments to obtain data from human subjects which would allow us to model the core descriptors. The purpose of this is to achieve a fuzzy-based knowledge base built exclusively on empirical data.

In 3.1) a thorough study of which operators should be used to combine the primitive fuzzy definitions should be made.
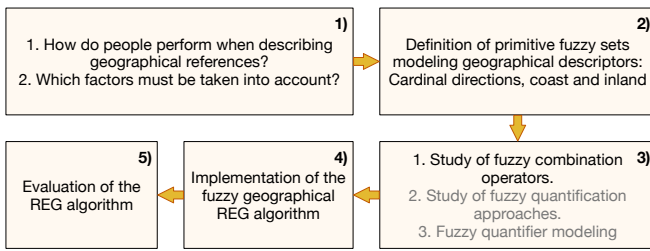
Fig. 2.  Diagram of the proposed methodology for addressing the geographical REG problem using a fuzzy approach.

This is inspired by the pioneering work of Zimmermann and Zysno [22], where the actual need of defining compensatory operators (defined as a new type of intermediate operators between the usual fuzzy conjunctions and disjunctions) was shown through a real experiment. Experimentation results showed that the compensatory operators more closely resembled the human way for knowledge representation and reasoning. Other similar approaches such as the OWA operators by Yager should also be taken into account [23].

An exploration of the use of fuzzy quantified sentences for this specific problem is also considered in 3.2) and 3.3), although this would also require to study the behavior of fuzzy quantification models in this particular case, as well as to define fuzzy quantifiers using a similar approach as in 2). This is motivated by the use of quantified expressions in the REG stage of the RoadSafe system [24], where target geographical referring expressions such as "in some far southern and southwestern places" were generated using crisp quantifiers.

Finally, 4) and 5) would involve the implementation of the GREs generation algorithm and its evaluation against classical approaches. Its subsequent integration into a newly developed or existing NLG system, such as RoadSafe, would also require a proper evaluation of the system as a whole.

With this methodology, which can easily be generalized for any other potential applications of fuzzy logic in NLG, our aim is to bring together the empirical techniques developed in the NLG community (corpus analysis, psycholinguistic experiments) with the theoretical models developed in the fuzzy logic and linguistic description of data community. The hope is that bringing these together will benefit both fields generally, and specifically will make it easier to develop algorithms for choosing qualitative descriptors which have a solid basis in both theory and empirical data.

Following the proposed ideas, we have run as part of step 1) an exploratory experiment in order to study how human subjects perform when referring to geographical descriptors.

## IV.  DESCRIPTION OF THE EXPERIMENT

### A.  Context

It is accurate to state that in most of the literature in linguistic description of data where practical use cases are presented, every fuzzy linguistic label or partition is usually defined by the authors, regardless of the application domain. This has not been an important issue in general (in many cases because the intent is to show how theory could be applied in

practical cases, and in others because evaluation by experts was performed and turned out positive, e.g. [11], [25]).

One simple and direct approach to modeling the primitive natural language descriptors in the problem of geographical referring expressions would be similar to the one proposed by Glöckner and Knoll in [26] for images. Cardinal directions can be defined as fuzzy sets on the latitude and longitude numeric domains, whereas descriptors such as "coastal" or "inland" could be defined based on the minimum distance from a point to the coast. For instance, Fig. 3 shows how the expression "southeast" could be obtained by the minimum operator by combining primitive self-defined fuzzy expressions "south" and "east".

However, in order to aim at applying such techniques in NLG systems that perform in real environments where expert knowledge is essential, such as RoadSafe [16], proper empirical acquisition of human knowledge should be performed. In this regard, for the particular case of generating GREs using fuzzy techniques, our first step in this direction has consisted in running an exploratory experiment to check:

- How human subjects perform when referring to geographical expressions.
- Additionally, factors that may influence the subjects' answers and thus should be taken into account in a proper fuzzy definition of the geographical descriptors. These include:
1) Subjects' prior knowledge of the geography.
2) Influence of the shape of a region in the perception of the directional descriptors.
3) Influence of the proximity of higher level terrain (e.g. mountain ranges) in the perception of the "coastal" descriptor.

Other factors unrelated to the morphology of the terrain have also been considered, such as proximity to population centers or the map scale, but were discarded in this
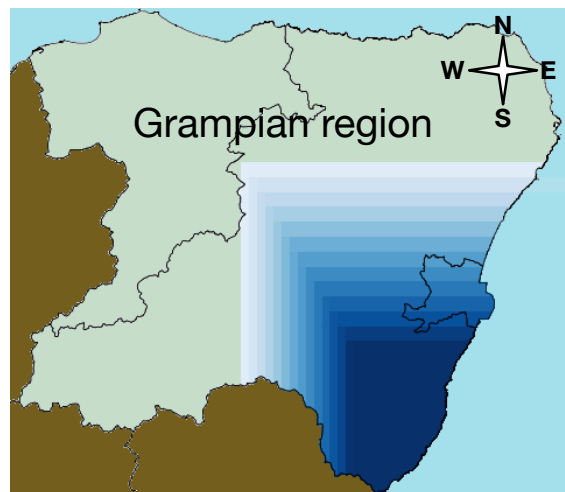


$$\mu_{SOUTHEAST} = \mu_{SOUTH} \cap \mu_{EAST}$$

Fig. 3.  Graphical representation of the fuzzy expression "southeast" for the Grampian region in Scotland.

case for simplicity reasons.

### B. Experiment design

The experiment was designed to be as simple as possible, without involving participants in complex tasks, which could be completed through a web-based interface in a reasonable amount of time (10-20 minutes). In a within-subjects design, participants from two different groups saw 66 different combinations of map locations and geographical referring expressions. They were asked to move a slider according to how appropriate they thought the given expression was for each given map location. Breaks were introduced every 20 questions (3 breaks) in order to partly alleviate the task repetitiveness and improve the user experience.

*1) Materials:* A single map was prepared for the Grampian region in northeastern UK, taken from Google Earth. A semi-transparent layer with the shape of this region was drawn on the map to highlight its boundaries, relevant mountainous locations were included in the form of green icon markers and a compass rose was also included on the top right corner of the map.

Using the previous map as basis, eleven different stimuli were created. Each stimulus included a single location which was indicated on the base map using a red pin (see Fig. 4). Locations used to create the stimuli were chosen intuitively as points of interest which fulfilled the geographic descriptors described below with regard to their relation to the coast and cardinal direction.

Six different geographical referring expressions (GREs) were considered for this experiment. The used GREs reflected whether the point was on the coast or inland (coast/inland descriptors), and whether the point was on the east, west, north and south (cardinal direction descriptors): "Inland Grampian", "Coastal Grampian", "Northern Grampian", "Southern Grampian", "Eastern Grampian" and "Western Grampian".

The experiment was composed of 66 questions. This number is a result of all the possible combinations between the six GREs and the eleven stimuli. For instance, the screenshot of the experiment interface in Fig. 4 shows a specific question which asks the subject to rate the appropriateness of the GRE "Coastal Grampian" for a given stimuli.

*2) Participants:* Participants were recruited in two countries: UK (University of Aberdeen, located in the Grampian region used for the maps in the stimuli) and Spain (University of Santiago de Compostela) through university mailing lists. The former represent participants who are familiar with the region (experts), and the latter represent those that are not (novices). The idea of using two different control groups according to their prior geographical knowledge of the map used in the experiment was inspired by a psycholinguistic experiment ran by Isaacs and Clark in [27], where the object of study was how experts and novices assess each other's expertise in a certain discussion topic and accommodate to their differences.
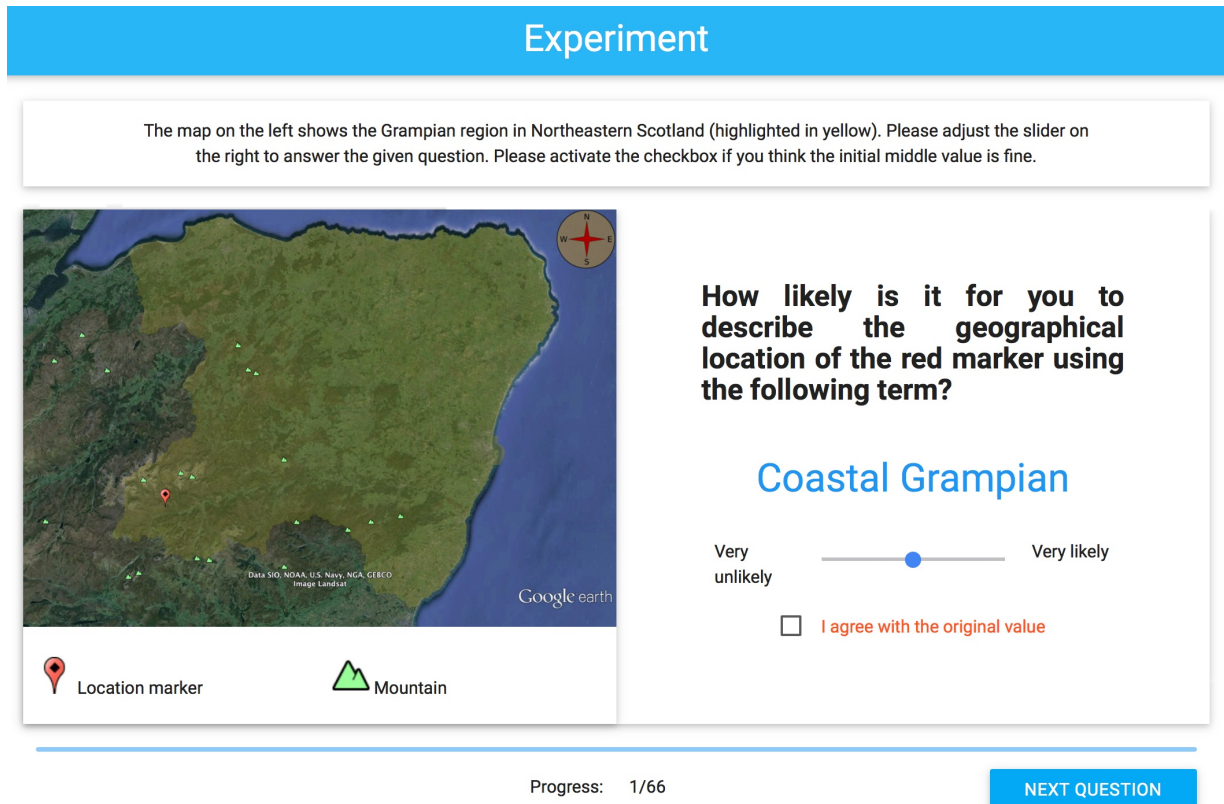


Fig. 4. Screenshot of the experiment interface.

*3) Procedure:* Participants were asked first about their origin, so that they could be placed in their corresponding group (UK/Spain). Once answered, subjects were shown a map with a point, and given a referring expression. They were then asked "How likely is it for you to describe the geographical location of the red marker using the following term?". Participants indicated the value (0=very unlikely, 100=very likely) of the likelihood by adjusting the position of a slider control, which was initialized with its middle value every time the subject advanced to the next question. The values the slider could take were numbers in the range [0,100]. Figure 4 shows the experiment interface, including the sort of map participants saw.

It must be noted that subjects were not aware of the value of the slider as it was not displayed. The purpose of this was to avoid participants sticking to certain values from one question to another, so that they would rely just on their intuition. Furthermore, in those cases where the slider was not moved, subjects were asked to activate a checkbox to confirm that they consciously agreed with the middle value as their answer.

Participants had to answer 66 randomized questions, one for each possible combination of referring expressions and stimuli. Afterwards, they could optionally send free-text comments about any aspect of the experiment. These have proven to be invaluable for shedding light at some of the results.

*4) Hypotheses:* This study is highly exploratory in nature, but it is driven by the following main qualitative hypotheses;

1) Participants will use a wide distribution of values (not only 0 and 100).
2) The degree of expertise will influence the likelihood of using a wider range of values. Non-experts will use a wider range of values.

Additionally, two secondary hypotheses regarding geographical features have been considered;

3) The shape of the region will influence the perception of cardinal directions.

4) The proximity of mountainous zones will influence the perception of the "coast" descriptor.

## C. Results

We received 27 responses in total (1782 questions were answered). These were distributed in 16 subjects for the expert group and 11 subjects for the novice group. Global results show there is a noticeable difference in the distribution of answer values between the two control groups (see Fig. 5).

*1) People use a wide range of values (H1):* It is true that both groups show a certain spread regarding the distribution of the answer values, but results differ greatly between them. For instance, Fig. 5 shows that more than 60% of the answer values given by the expert group are placed on the lowest extreme of likelihood, while less than 40% are spread within the remaining value range ([10,100]). However, in the case of the novice group these percentages are much more even (near 40% for "very unlikely", near 20% for "very likely" and thus around 40% for the rest of values).

One factor which influences the presence of low likelihood values in both groups is that locations used for the stimuli were not homogeneously spread across the map, as they were partly focused on the two secondary hypotheses. This means that some descriptors such as "coastal","inland" or "north" had more representation than others and thus, in general, answers around high likelihood have a higher presence in those specific cases. Likewise, this also means that very low likelihood values having a higher presence globally were to be expected to a certain extent. However, this trend was clearly surpassed in the case of the expert group and required further exploration, which is discussed in *H2*.

*2) Level of expertise affects which values are used (H2):* Our initial hypothesis about the performance of both groups considered that there would be slight differences about how values would spread: experts would be more clear about how to relate each location to the geographical referring expressions and novices would be more erratic and provide more spread
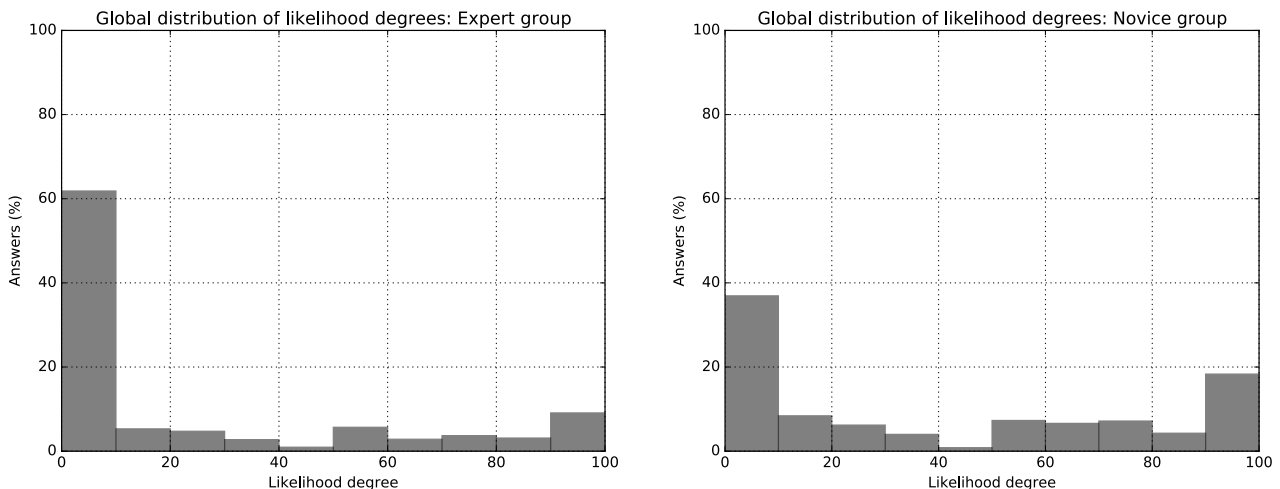


Fig. 5. Global percentage distributions of the likelihood answer values given by both expert and novice groups.

answers. While this has turned out to be true, the answers from the expert group had a higher proportion of responses with the value 0 than for the novice group, implying that factors we had not considered were influencing their answers.

Free-text comments made by expert participants shed light on this matter:

- *...I felt all were unlikely. Maybe an example of the full description would have helped. I would normally use northern to indicate a region, not an individual point. However if it was city A is in northern Grampian then yes, would become easier to do.*
- *I have answered 'Very Unlikely' to all questions as I don't use the term 'Grampian'. I will use 'inland' or 'on the coast'. I'm most likely to describe places by where they are near, such as 'out Alford way' or 'up by Fraserburgh'.*
- *As a local I don't use Grampian for describing where places are within this area. I would say the city, or the shire, or Moray, or Banff etc. so I had to disagree with all of the questions.*

Many of the participants in the expert group answered all of the questions as "very unlikely" due to not using geographical descriptors to refer to locations, but other kind of descriptors such as proper names. This issue is highly related to research in previous work [18], where it is shown that people mostly use named entities to refer to locations of geographic scale. Nonetheless, the formulation of the question "How likely is it for you to describe the geographical location of the red marker using the following term?" appears to have influenced expert participants in a misleading way.

Some of the novices contacted the researchers asking for clarification before completing the experiment, but no comments were received regarding the previous issue (which was to be expected since they are not familiar with the map and locations they were shown). Thus, it is safe to assume that results for the novice group depend exclusively on the geographical intuition of the participants. Given that the results of the expert group are biased by the aforementioned issue, we have checked *H3* and *H4* based solely on the results of the non-expert group.

*3) Region shape influences perception of directions (H3):* As part of the experiment, we have explored additional factors which, in our opinion, might influence how human subjects perceive certain geographical descriptors. For this, a few stimuli were specifically prepared to study *H3* and *H4*.

With *H3*, our aim was to check whether the region shape could be an important factor for participants when referring to cardinal descriptors ("Eastern Grampian" in this case). Thus, two stimuli (Stimulus 3 and 13) were prepared in such a way that they showed two locations with similar longitude values, but were located at different distances from the eastern coast (see Fig. 6). Results show that, in general, participants considered the location in Stimulus 3 more appropriate for the referring expression "Eastern Grampian" (likelihood average: 58, median: 71) than the one shown in Stimulus 13 (likelihood average: 39, median: 60). This seems to indicate there was an influence on how subjects perceived the "east" descriptor, in such a way that the location farther from the shapeline side related to the descriptor (Stimulus 13) got lower likelihood values.

*4) Mountain proximity influences perception of the "coast" descriptor (H4):* The existence of high level terrain close to coast zones was a factor we considered for influencing how human subjects perceive the "coast" descriptor. In a similar fashion to *H3*, two stimuli were prepared in such a way that both location markers were placed at a similar (relatively close) distance from the coast, but one of them was on a mountain (Stimulus 1) and the other one (Stimulus 2) was on low terrain (see Fig. 7). In this case, participants gave Stimulus 2 a higher likelihood for the GRE "Coastal Grampian" (likelihood average: 54.36, median: 70) than Stimulus 1 (likelihood average: 35.45, median: 32). This indicates that the proximity of a mountain influenced how subjects perceived the "coast" descriptor. This influence translates into lower likelihood degrees for the given descriptor.
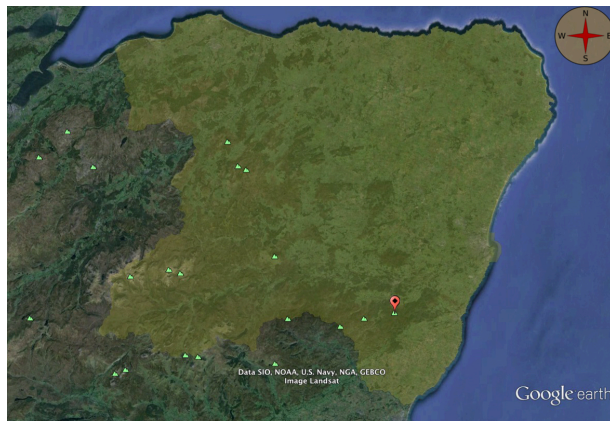
### D. Discussion

Although the experiment we have described is highly exploratory in nature (with a small sample size) and the answers



Stimulus 3 (closer to the eastern coast)    Stimulus 13 (farther from the eastern coast)

Fig. 6.   Stimuli 3 and 13, used for checking *H3*.

Stimulus 1 (on a mountain)  Stimulus 2 (on low terrain)

Fig. 7. Stimuli 1 and 2, used for checking *H4*.

from the expert group were likely to be influenced by how the experiment question was formulated, global results show that the degree of appropriateness matching geographical descriptors and map locations spreads across the whole likelihood value range (especially in the case of the novice group, see Fig. 5). Participants perceived a gradual relationship between the different locations used in the stimuli and the geographical referring expressions shown in the questions, which opens up the possibility of modeling such graduality through the use of fuzzy sets.

The focus of this exploratory research is to investigate how people use geographic descriptors in specific cases, and which factors influence their answers. The aim of this investigation is to understand the problems and complexities of GREs (Step 1 of the methodology proposed in Sec. III). However, a larger experiment with a modified design would be needed in order to complete the current step. Based on the experience here described it is feasible to point out which changes should be considered for a forthcoming experiment before advancing to Step 2:

- Providing a clearer question which does not lead subjects to misunderstanding or confusion.
- Aiming at a larger set of internationally spread experiment subjects ($\approx$100 participants at least). Although it should still be controlled, no explicit group distinction would be made regarding the familiarity or expertise of the subjects about the geographical region used in the experiment.
- Using a set of stimuli locations which covers the whole region of interest and represents geographical descriptors in a balanced way.
- Utilizing a slider-based interface may be tedious for participants when there is a large number of answers to supply. In an experiment where the number of answers is expected to be high, a set of checkboxes could be used to allow subjects to crisply relate one or more descriptors to the stimuli.

The results obtained in this improved experimental setting would allow us to aim at the task of modeling a set of fuzzy descriptors (Step 2). A voting model [28] which aggregates the obtained answers for each representative location could be

applied to obtain a distribution of fulfillment degrees for the set of geographical descriptors. Based on such distributions, the fulfillment degree of a given descriptor for any location in the region could be calculated based, for example, on an interpolation of the distributions of the nearest representative locations.

Using the model obtained in Step 2, Step 3 would involve studying the behavior of several fuzzy aggregation operators and quantification models in the task of combining descriptors for building more complex expressions (e.g. "southwest inland", "northwestern coast", "in some southwestern places"). This would allow to determine the most human consistent approach to the problem of generating GREs, which would then be implemented in Step 4 and evaluated in Step 5.

Additionally, more factors could also be considered as part of this research line:

- The geographical referring expressions used in this experiment here presented were considered separately, but these may be seen as natural antonyms ("coast"/"inland", "west"/"east", "north"/"south"). In a coherent model (and coherence is hard to achieve, even when knowledge is provided by experts from the same domain [29]) such kind of properties and relationships will play an important role in assuring the coherence of the GREs the model will generate.
- The same experiment could be rerun using different map scales to determine if there are significative differences among the obtained results.
- Although the current focus of this research is to establish a good methodology to build and use fuzzy models of geographical descriptors for specific problems, how to generalize the descriptors obtained from the experimental data is also a challenging problem which can be considered in the future.
- Other type of descriptors which, for instance, relate different regions (e.g. "the region to the south of Canada", "to the west of Germany") could also be explored.

## V. Conclusions

We have discussed the interest of using fuzzy set techniques to address the NLG problem of generating geographical referring expressions. Motivated by this interest, we have proposed a methodology to approach this problem and, as part of it, have run an exploratory experiment to study how people associate geographical locations in a map with different geographical descriptors. Although we partly failed at providing subjects with a clear description of the experiment, which may have resulted in biased answers from the expert group, results show that participants related locations and descriptors in a gradual way.

As future work, we intend to run another experiment based on the modifications described in the discussion in Sec. IV-D and perform statistical analysis to complete Step 1 of the methodology proposed in Sec. III. Afterwards, we will create a fuzzy model for primitive geographical descriptors and explore the use of fuzzy operators and techniques which will allow us to generate the target geographical referring expressions.

## References

[1] K. van Deemter, "Generating referring expressions that involve gradable properties," *Comput. Linguist.*, vol. 32, no. 2, pp. 195–222, Jun. 2006.

[2] ——, "Utility and language generation: The case of vagueness," *Journal of Philosophical Logic*, vol. 38, no. 6, pp. 607–632, 2009.

[3] R. Power and S. Williams, "Generating numerical approximations," *Comput. Linguist.*, vol. 38, no. 1, pp. 113–134, Mar. 2012.

[4] F. Portet and A. Gatt, "Towards a possibility-theoretic approach to uncertainty in medical data interpretation for text generation," in *Knowledge Representation for Health-Care. Data, Processes and Guidelines*, ser. Lecture Notes in Computer Science, D. Riaño, A. ten Teije, S. Miksch, and M. Peleg, Eds. Springer Berlin Heidelberg, 2010, vol. 5943, pp. 155–168.

[5] J. Kacprzyk, A. Wilbik, and S. Zadrożny, "Linguistic summarization of time series under different granulation of describing features," in *Rough Sets and Intelligent Systems Paradigms*, ser. Lecture Notes in Computer Science, M. Kryszkiewicz, J. Peters, H. Rybinski, and A. Skowron, Eds. Springer Berlin Heidelberg, 2007, vol. 4585, pp. 230–240.

[6] J. Kacprzyk and S. Zadrozny, "Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries, and natural-language generation," *Fuzzy Systems, IEEE Transactions on*, vol. 18, no. 3, pp. 461–472, June 2010.

[7] J. Moreno-Garcia, L. Rodriguez-Benitez, J. Giralt, and E. del Castillo, "The generation of qualitative descriptions of multivariate time series using fuzzy logic," *Applied Soft Computing*, vol. 23, pp. 546 – 555, 2014.

[8] N. Marín and D. Sánchez, "On generating linguistic descriptions of time series," *Fuzzy Sets and Systems*, vol. 285, pp. 6 – 30, 2016, special Issue on Linguistic Description of Time Series.

[9] A. Ramos-Soto, A. Bugarín, and S. Barro, "On the role of linguistic descriptions of data in the building of natural language generation systems," *Fuzzy Sets and Systems*, vol. 285, pp. 31–51, 2016.

[10] A. Bugarin, N. Marin, D. Sanchez, and G. Trivino, "Aspects of quality evaluation in linguistic descriptions of data," in *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, Aug 2015, pp. 1–8.

[11] A. Ramos-Soto, A. Bugarín, S. Barro, and J. Taboada, "Linguistic descriptions for automatic generation of textual short-term weather forecasts on real prediction data," *Fuzzy Systems, IEEE Transactions on*, vol. 23, no. 1, pp. 44 – 57, 2015.

[12] A. Gatt and F. Portet, "Multilingual generation of uncertain temporal expressions from data: A study of a possibilistic formalism and its consistency with human subjective evaluations," *Fuzzy Sets and Systems*, vol. 285, pp. 73 – 93, 2016, special Issue on Linguistic Description of Time Series.

[13] E. Reiter and R. Dale, *Building Natural Language Generation Systems*. Cambridge University Press, 2000.

[14] R. Dale and E. Reiter, "Computational interpretations of the gricean maxims in the generation of referring expressions," *Cognitive Science*, vol. 19, no. 2, pp. 233 – 263, 1995.

[15] E. Krahmer and K. van Deemter, "Computational generation of referring expressions: A survey," *Comput. Linguist.*, vol. 38, no. 1, pp. 173–218, Mar. 2012.

[16] R. Turner, S. Sripada, E. Reiter, and I. P. Davy, "Using spatial reference frames to generate grounded textual summaries of georeferenced data," in *Proceedings of the Fifth International Natural Language Generation Conference*, ser. INLG '08. Stroudsburg, PA, USA: Association for Computational Linguistics, 2008, pp. 16–24.

[17] ——, "Selecting the content of textual descriptions of geographically located events in spatio-temporal weather data," *Applications and Innovations in Intelligent Systems*, vol. XV, pp. 75–88, 2007.

[18] R. de Oliveira, Y. Sripada, and E. Reiter, *Proceedings of the 15th European Workshop on Natural Language Generation (ENLG)*. Association for Computational Linguistics, 2015, ch. Designing an Algorithm for Generating Named Spatial References, pp. 127–135.

[19] L. A. Zadeh, "Fuzzy logic = computing with words," *Fuzzy Systems, IEEE Transactions on*, vol. 4, no. 2, pp. 103–111, 1996.

[20] R. R. Yager, "A new approach to the summarization of data," *Information Sciences*, vol. 28, no. 1, pp. 69 – 86, 1982.

[21] R. R. Yager, K. M. Ford, and A. J. Cañas, "An approach to the linguistic summarization of data," in *Uncertainty in Knowledge Bases, 3rd International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 90, Paris, France, July 2-6, 1990, Proceedings*, ser. Lecture Notes in Computer Science, vol. 521. Springer, 1990, pp. 456–468.

[22] H.-J. Zimmermann and P. Zysno, "Latent connectives in human decision making," *Fuzzy Sets and Systems*, vol. 4, no. 1, pp. 37 – 51, 1980.

[23] R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, no. 1, pp. 183–190, Jan 1988.

[24] R. Turner, S. Sripada, and E. Reiter, "Generating approximate geographic descriptions," in *Empirical Methods in Natural Language Generation*, ser. Lecture Notes in Computer Science, E. Krahmer and M. Theune, Eds. Springer Berlin Heidelberg, 2010, vol. 5790, pp. 121–140.

[25] L. Eciolaza, M. Pereira-Fariña, and G. Trivino, "Automatic linguistic reporting in driving simulation environments," *Applied Soft Computing*, vol. 13, no. 9, pp. 3956 – 3967, 2013.

[26] I. Glöckner, *Fuzzy Quantifiers: A Computational Theory*. Springer, 2006.

[27] E. A. Isaacs and H. H. Clark, "References in conversation between experts and novices," *Journal of Experimental Psychology: General*, vol. 116, no. 1, pp. 26–37, 1987.

[28] J. Baldwin, J. Lawry, and T. Martin, "A mass assignment theory of the probability of fuzzy events," *Fuzzy Sets and Systems*, vol. 83, no. 3, pp. 353 – 367, 1996.

[29] E. Reiter, S. Sripada, J. Hunter, and I. Davy, "Choosing words in computer-generated weather forecasts," *Artificial Intelligence*, vol. 167, pp. 137–169, 2005.