# Modeling the User's Preferred Perceptual System by Analyzing the User's Wording

Gudrun Kellner

Vienna University of Technology, Institute of Software Technology and Interactive Systems,
E-Commerce Group, Favoritenstr. 9-11/188, 1040 Vienna, Austria
kellner@ec.tuwien.ac.at

**Abstract.** A new method for user modeling is examined, assuming that the information on the individual perceptual preference is encoded in every text the user is producing; the choice of expression may be considered as the cue.

**Keywords:** User Modeling, Perception, Language use

## 1      Introduction

One of the great challenges in User Modeling is still the question on how to obtain personal information about the user. Of course, the most direct way is to ask him or her. Nevertheless, this option is not much appreciated by the user: it takes time and effort to fill out questionnaires and often gives the user a feeling of being kept from his or her original goal. On the other hand it is not sure that the user is always conscious of the information needed by the system, especially when it gets to complex aspects of his or her personality or his or her preferences. It would hence be very helpful to find ways to get the information of interest without the need of asking the user first.

This paper makes an attempt to use the vast information on users provided by their own postings in forums as indication of their perceptual preferences. Such implicitly gained knowledge about the user's perceptual preferences would be of great help for presenting information to the user in a more comprehensive way and furthermore may give cues on preferences in other fields. The concept that promises to deliver such additional knowledge on the user is taken from the research field of didactics. It assumes that every person has one or two preferred sensory systems for perceiving information. Those sensory systems are built as equivalent to the five senses and can be split up into visual, auditory, kinesthetic (as compound for tactile and interior sensation), olfactory and gustatory perceptual information. The idea of perceptual preferences is, like other description models, a construct for explaining cognitive processes, so one can hardly get direct and measurable input and therefore rather needs to filter implicit information for some indication. Thus, another theory is taken into consideration: The psycholinguistic idea of getting information on the interlocutor's personality by analyzing his or her choice of expression. A combination

2      **Gudrun Kellner**

of both leads to a new method for identifying perceptual preferences via text published in the internet by the detection of cue words.

For establishing that method on a solid base, three fields of research need to be examined: the field of learning types, the field of language and personality, and the methods used in the field of user modeling and adaptation, in particular the findings with respect to the first two fields.

*Learning types.* One focus in didactical research is on presentation and perception of information. An application of this concept for the field of language learning was published by *Reid* [1]. In her model on learning styles she differentiates between four basic perceptual learning channels, distinguishing visual, auditory, kinesthetic and tactile modalities, and presents a questionnaire focusing on the learner's perceptual preference and his or her preference of individual vs. group learning.

*Language and personality.* Talking as well as writing is not only a way of showing the personal style of expression, but also a way of describing the personal point of view which is always guided by the individual focus of interest. An interesting insight in the relationship of language use and personality is given by *Nowson* [2]. He shows that both gender and personality are projected by language in blogs. Another field of research related to this topic is on dimensions of subjectivity. *Chen* [3] mentions the dimensions non-objectivity, uncertainty, vagueness, non-objective measurability, imprecision and ambiguity. *Wiebe et al.* [4] show that unique words are subjective more often than expected and that unique words are valuable clues to subjectivity. Perceptual preference can be regarded as another dimension of subjectivity, giving cues on the individual point of view. Considering the perceptual preference as another specification of the personal point of view, those preferences are in all likelihood also expressed in the individual way of telling a story.

*Methods used in the field of user modeling and adaptation.* A related field of research that also plays an important role in the human-computer interaction is the research on emotional states, showing one significant analogy: in that field it is not possible either to collect data in the direct way by asking the user. First starting with the collection of physiological data like heart rate monitoring, skin conductance, and measurement of brain activities, those methods turned out to be applicable in test settings, but not in situations of everyday use for reasons of realization problems caused by infrastructural and cost factors as well as for the need of expertise for using those instruments. Hence other ways to collect that information were tested. Two research projects that try to use language as a source for information are one on doubt recognition by *Carberry et al.* [5] and one on affect detection by *D'Mello et al.* [6]. While the first focuses on the lexical and semantic aspects of expressing doubt, the second one uses a multimodal approach, taking also into account facial and acoustic-prosodic features and combining several classification methods.

Our project consists of three phases, where each one is reflected in a section of this paper: The first one concentrates on the creation of a corpus of sensory expressions, the second one focuses on the development of an adequate analyzing tool for filtering

those expressions, the third one explores the relation of the user's wording, his or her perceptual preference, and his or her fields of interest.

## 2      The vocabulary of sensory expressions

For identifying the individual perceptual preference, we decided to use lexical keywords. In most languages exists a huge number of words that are linked directly to one sensory system. The related work in that field contains a lot of publications discussing a usually very narrow aspect of sensory expressions. Most of them concentrate on only one sensory system, often only on one lexical category, some of them discussing only two to five words and their composites into depth. [7, 8, 9, 10, 11] The only corpus of our knowledge in the German language which includes words and expressions of all sensory classes and of all different lexical categories is presented by *Kellner* [12]. She built a partial corpus of sensory expressions with the input of 13 different resources, containing more than 840 different entries with around 200 words and expressions per perceptual class.

As our goal is to identify perceptual preference only by lexical cues, we are trying to gain the most complete corpus of sensory expressions possible. Therefore we initiated a cooperative project with the Institute of German Studies at the University of Vienna, where the entire German lexicon is analyzed, filtered, and linked to a sensory system on the following criteria:

- If the word is expressing sensory activity, the word belongs to the sensory class which perceives an input (like "watch" as visual verb, "taste" as gustatory verb).
- If the word is describing an object, an action, or an attribute which can be perceived by only one sensory system, the word belongs to the sensory system that needs to be active (like "light" as visual noun or adjective, "talk" as an auditory verb as it can only be perceived by the auditory system, "hard" as kinesthetic adjective in both senses: either explaining a surface or a feeling).
- If the word is the written version of a sound (ling. *onomatopoeia*), it belongs to the auditory system (like "haha" for the sound of somebody laughing or "woof" for the sound of a dog barking).

The outcome of the project "Convergences of perceptual and linguistic spheres: The German vocabulary of perception and its interdisciplinary relevance" is a complete list of German sensory expressions that can be considered as cue words for perceptual preference which can be used as basic information for the analyzing tool.

## 3      The analyzing tool

The second big part of our research consists of the conception and development of a software tool that is able to filter the sensory expressions included in the corpus from any text.

4        **Gudrun Kellner**

One of the challenges in that part will be the development of an adequate search algorithm. Being a compounding language, the German language allows the generation of complex words by merging together simpler ones. Therefore a special treatment is needed for finding all possible compound words "hiding" one of the words listed in the corpus. For example, when searching for the signal word "rot" [red], the output should also include words like "Morgenrot" [dawn] and "rotgrün" [red and green] containing the same semantic root, but not words like "Brot" [bread] or "Schrot" [bruised grain] which include the same sequence of characters, but have no semantic link to the searched word. An examination of actual methods in the field of computational linguistics on stemming and its alternatives will lead to a description model of word grammar. According to the chosen word description model has to be decided whether it could be helpful to include parsing methods as well. Particular attention shall be given to the format of the searched text (forum entries, eventually blogs, both including HTML meta-information that needs to be filtered). The final search algorithm will be a result of all those reflections and will be an executable step-by-step description of how to find cue words and compound words containing a cue word in a text.

## 4        Sensory expressions in forums: Do they make *sense*?

The most crucial point in our research project will be the design of the online-survey. There are several hypotheses that shall be verified or falsified. These hypotheses are:

H1   *Most users have one or two preferred sensory systems they are using significantly more often than the others, regarding their choice of expressions.*

H2   *The choice of expressions, regarding the sensory systems, may also rely to the topic a user is writing about.*

The first hypothesis is the basic one. It shall bridge the didactic concept of perceptual preference with the individual use of language as a possibility to express personality and show that the catenation of both is a valid link-up. Nevertheless it is most likely that the choice of expression is not solely influenced by the perceptual preference, but also depends on the topic a user is writing about; e.g. when discussing about photography, there will be a higher use of visual vocabulary than in a conversation about music. That fact may be of particular importance if the choice of topic correlates significantly with the user's fields of interest. A list with fields of interest of reasonable size is used by [13].

H3   *In the list of "fields of interest" there are some fields of interest seeming to be more related to one sensory system than to the others.[1]*

---

[1] This may e.g. be the case for "cinema" as a rather visual and "sports" as a rather kinesthetic domain.

**Modeling the User's Preferred Perceptual System by Analyzing the User's Wording**    5

Tourism as the field of application for our research brings one more hypothesis related to the special aspects of this research domain. It shall be examined whether there exists a relation between the perceptual preference and the choice of holiday activities according to the travel personality types presented in [14] and tourist types explored in [15] for the use in travel recommender systems [16].

*H4    The preferred holiday activities may rely to the preferred sensory system.*

As the interest of those hypotheses is quite broad, the online-survey has to cover several different points. It shall be composed by the following parts:

- A questionnaire concerning the "fields of interest" as a check list with the possibility to choose as many as wanted, and a listing of the three fields with the highest importance to the user.
- A list of possible holiday activities as a check list (including a number of activities that are expected to rely to one sensory system), using the same system as above.
- Still in discussion is a part on either a test or a self estimation about the perceptual type. If this should be a test, one could use the famous test of *Reid* on learning types [1], using the questions concerning the perceptual types. If this should be a self estimation, it could be realized by a question like "Would you consider yourself being rather a visual/auditory/kinesthetic/olfactory/gustatory type?" (including explanations for each perceptual type). At least one of those two options shall be realized due to the fact that employing multiple measures is the standard criterion for establishing convergent validity [17].

The second step is to check all the forum entries published by the user for the usage of perceptual expressions. For reasons of significance, the samples that under-run a minimum of altogether length of text produced by the user shall be eliminated. Only the data of users fulfilling both criteria (having published enough text and having filled out the online questionnaire) will be evaluated. For raising the motivation to do so, there will be a price draw among the participants of the study.

The hypotheses testing shall be realized by (*H1*) analyzing the use of sensory expressions occurring in the user's forum entries, (*H2*) comparing the topics the user is writing about in his forum entries and the used sensory expressions, (*H3*) comparing the use of sensory expressions to the chosen fields of interests, and (*H4*) comparing the use of sensory expressions and the chosen fields of interests to the choice of holiday activities.

## 5    Conclusion

If the expected relation between the preferred perceptual system and the used sensory expressions can be verified, the tested method could be a great resource for user modeling: Information on the user's perceptual preference could then be obtained without the need of questionnaires, but by using our analyzing tool instead. If the relation between the perceptual preference and the choice of holiday activities can be proved, a new concept of recommendation systems for the tourism domain is given.

6        **Gudrun Kellner**

The results – if showing the expected relations – will not only be applicable in the tourism domain, but probably also in several other fields of electronic commerce. Another possible field of use is the sector of e-learning: With the knowledge of the learner's preferred perceptual system, content could be presented in a (sensory) way providing easier understanding for the learner.

# References

1. Reid, J.M.: The Learning Style Preferences of ESL Students. In: TESOL Quarterly, 21/1, pp. 87–111 (1987)
2. Nowson, Scott: The Language of Weblogs: A study of genre and individual differences. Doctoral Thesis. University of Edinburgh (2006)
3. Chen, W.: Dimensions of Subjectivity in Natural Language. In: Proceedings of ACL-08: HLT, Short Papers (Companion Volume), pp. 13–16 (2008)
4. Wiebe, J., Wilson, T., Bruce, R., Bell, M., Martin, M.: Learning Subjective Language. In: Computational Linguistics, V. 30, No. 3, pp. 277–308 (2004)
5. Carberry, S., Schroeder, L., Lambert, L.: Toward recognizing and conveying an attitude of doubt via natural language. In: Applied Artificial Intelligence, 16, pp. 495–517 (2002)
6. D'Mello, S.K., Craig, S.D., Witherspoon, A., McDaniel, B., Graesser, A.: Automatic detection of learner's affect from conversational cues. In: User Model. User-Adapt. Interact, pp. 45–80 (2008)
7. Robering, K.: Die deutschen Verben des Sehens. Eine semantische Analyse. Kümmerle, Göpping (1985)
8. Schepping, M.-T.: Kontrastive semantische Analyse von Verben des Visuellen im Französischen und Deutschen. Narr, Tübingen (1982)
9. Vliegen, M.: Verben der auditiven Wahrnehmung im Deutschen. Eine semantisch-syntaktische Analyse. Narr, Tübingen (1988)
10. Fricke, E.: Die Verben des Riechens im Deutschen und Englischen. Eine kontrastive semantische Analyse. KIT-Report 136, TU Berlin (1996)
11. Linke, A., Nuessli, J. (eds.): Semantik der Sinne. Proceedings zur Tagung. 18. und 19. Juni 2004, Universität Zürich. Labor für Lebensmittelchemie und -technologie, ETH Zürich (2005)
12. Kellner, G.: Wege der Kommunikationsoptimierung. Anwendung von NLP im Bereich der Künstlichen Intelligenz. Master Thesis, University of Vienna (2006)
13. List of fields of interest collected by gmx.net, http://www.gmx.net
14. Gretzel, U., Mitsche, N., Hwang, Y.-H., & Fesenmaier, D.R.: Tell me who you are and I will tell you where to go: Use of Travel Personalities in Destination Recommendation Systems. In: Information Technology and Tourism, 7(1), pp. 3–12 (2004)
15. Berger, H., Dittenbach, M., Denk, M., Merkl, D. and Pesenhofer, A.: Getting tourists quicker to the fun part: Photographs, tourist types, community beings, and their implications for a 3d e-tourism environment. Information Technology and Tourism, Vol. 9, No. 3–4, pp. 211–226 (2007)
16. Ricci, F., & Werthner, H.: Case based Reasoning for Travel Planning Recommendation. Information Technology and Tourism, 4(3–4), pp. 215–226 (2002)
17. Chin, D.: Empirical Evaluation of User Models and User-Adapted Systems. In: User Modeling and User-Adapted Interaction 11, pp. 181–194 (2001)