

uComp Language Quiz - A Game with a Purpose for Multilingual Language Resource Acquisition

Arno Scharl

MODUL University Vienna
Department of New Media Technology
Am Kahlenberg 1, Vienna, Austria
scharl@modul.ac.at

Michael Föls

Vienna Univ. of Economics & Business
RI for Computational Methods
Welthandelsplatz 1, Vienna, Austria
michael.foels@wu.ac.at

Abstract

This paper presents the *uComp Language Quiz*, an online application in the tradition of games with a purpose for language resource acquisition. It is based on the human computation framework developed within the uComp research project (www.ucomp.eu), which provides multi-channel deployment and social logins, a viral notification system, quality control mechanisms, and a CrowdFlower data interface to publish game elements as *Human Intelligence Tasks*.

1 Introduction

The Language Quiz, available at quiz.ucomp.eu, helps to acquire multilingual language resources for research purposes. A flexible and modular architecture to support multiple languages and task types were among the key goals that have guided the development efforts. Conceptual insights gathered from analyzing the results of the *Sentiment Quiz*, a game with a purpose to assess sentiment terms (Scharl et al., 2012), also influenced design decisions. Both the crowdsourcing engine and individual game components were continually refined based on the feedback from beta testers and early adopters. Special emphasis was placed on the interaction design - including the scoring and engagement mechanisms as well as the flexible support of different task types. Questions where the correct answer will never be known ("opinion poll"), for example, or questions where the correct answer will be determined at a future point in time ("prediction"). In the former case, the points are based on the mean answer of all players. In the latter case, the points are awarded ex post once the correct answer is known.

2 Workflow and Task Management

An *Application Programming Interface* (API) to send tasks to the game and retrieve results without further administration effort currently supports multiple choice and sentiment assessment tasks - on a term or a sentence level (the latter serving either stand-alone statements, or highlighted sentences in a context of a whole paragraph). When uploading the task data, the project owner can decide if the tasks should be sent (i) to the game, (ii) to CrowdFlower, or (iii) to a hybrid workflow that uses CrowdFlower to recruit game participants.

The mechanics of the game were also adapted to support the hybrid scenario. If a user is sent to the game via Crowdflower, he would receive 20 questions. Upon successful completion of the game rounds, the user will receive a code which he can enter in his Crowdflower task to receive payment. Since only the highest-rated Crowdflower users were being targeted, the quality of the results were correspondingly good. To filter out cheating players, each Crowdflower user has to answer three gold standard questions. In the case of a wrong answer on a test question, the user's answers will be deleted from the game. This ensures a consistent high quality of the obtained results.

The hybrid support of native gaming and paid Crowdflower evaluation leverages the benefits of both crowdsourcing categories depending on specific task requirements. Paid tasks also improve the user experience of the other game participants, since correct answers and related game point calculations are available sooner. Games with a purpose and mechanized labour have specific advantages and disadvantages in terms of speed, cost and quality of answers (Sabou et al., 2013). Using the uComp framework, task providers can make the decision on a per task basis.



Figure 1: Screenshot of the uComp Language Quiz

3 Quality Control

Cheating prevention is an essential element of crowdsourcing applications. If Crowdfunder is only used as a promotional tool to recruit players for a certain task, there are two systems in place to prevent cheating: (i) Users have to insert a code in the Crowdfunder submission form, which they will receive after having answered 20 questions in the game. This code is a complex set of characters, which cannot be guessed; (ii) Crowdfunder allows to specify a minimum time that it usually takes to complete a task (filtering out users who do not provide a genuine answer); e.g. if the minimum time is set to two minutes, but the answers is received within 10 seconds, it can be assumed that the user has not actually tried to solve the problem.

If Crowdfunder is used to generate the actual results (i.e. the uComp API only serving as an interface between task owner and crowdsourcing marketplace), each player has to correctly answer a certain number of gold standard questions as part of the selected batch of tasks. This identifies users who attempt to cheat or those who lack the expertise for a specific task (e.g. a lack of language skills). Answers of players who fail one test question will be disregarded.

The additional complexity introduced by redirecting participants to the game interface suggests to only consider highly rated "Level 3" players to solve the tasks, except in time-critical situations. Those players tend to be more careful when solving a task, as they risk their Level 3 status if they fail too many test questions.

4 Engagement and User Statistics

Since its launch in October 2015, the *Language Quiz* attracted 2,688 unique users. 2,150 of these users created an account, and 2,002 became active players (1,916 of these players were contacted via the Crowdfunder campaign option). 959 active players submitted valid answers, while 1,043 failed the test questions. In total 65,021 valid answers were submitted to the game - 55,791 of those answers were paid via Crowdfunder, 9,230 where organic answers from players. In terms of language distribution, the five dominant languages of the result set were French (55.9%), English (24.4%), Russian (11.9%), Spanish (6.7%), and German (1.0%).

Acknowledgments

The presented research is part of the uComp project (www.ucomp.eu), funded by the Austrian Research Promotion Agency (FWF 1097-N23) in the CHIST-ERA ERA-NET program line.

References

- M. Sabou, A. Scharl, and M. Föls. 2013. Crowdsourced knowledge acquisition: Towards hybrid-genre workflows. *International Journal on Semantic Web and Information Systems*, 9(3):14–41.
- A. Scharl, M. Sabou, S. Gindl, W. Rafelsberger, and A. Weichselbraun. 2012. Leveraging the wisdom of the crowds for the acquisition of multilingual language resources. In Nicoletta Calzolari, editor, *8th Int'l Conference on Language Resources and Evaluation (LREC-2012)*, pages 379–383.