Borsa Parole – A Market for Linguistic Speculation

Fabian Kneissl and François Bry

Ludwig-Maximilian University of Munich, Germany

30 March 2012

Abstract

This article describes a novel approach to linguistic field research consisting in exploiting the self-regulation of a market for collecting data on language use. The market is conceived as an output-agreement game with a purpose called Borsa Parole. The agreement can be traded with by the players what makes it adjustable. Borsa Parole has been conceived and is deployed for a linguistic study on the divergence of Italian dialects and vernaculars.

Introduction

In traditional linguistic field research, studies are conducted by researchers, typically doctoral students, interviewing in the various areas of relevance on the language they use people of various backgrounds. For being meaningful, a large number of people must be interviewed, what makes linguistic field research a time consuming and expensive endeavour. As a consequence, mostly small scale studies are usually conducted. Eventually, the data gathered from such small scale studies are brought together, what might result in inconsistent data. Furthermore, the data being collected by researchers are rarely unbiased. Biases might affect how people are selected and how their answers are written down by the linguists conducting the field research [4, 13].

This article describes a new game with a purpose (GWAP) with code name "Borsa Parole", Italian for "Word Stock Exchange". A beta version of the game is available at `http://beta.metropolitalia.org/`. Borsa Parole has been specifically designed for overcoming the aforementioned drawbacks of traditional linguistic field research. Borsa Parole leaves it to players to both suggest vernacular or dialectal sentences and to estimate the geographical areas or social communities in which these sentences are used. Borsa Parole keeps the linguists out the data collection loop, thus both dramatically reducing the costs of field research and ensuring an unbiased data collection.

The gaming paradigm of Borsa Parole is that of a stock exchange for vernacular or dialectal sentences. Players can suggest sentences in a dialect or vernacular, assign them geographical areas where they think these sentences are commonly used and furthermore give an estimate of the

proportion of people, that is, players, that are likely to assign the same areas to these sentences. A player performs well if her sentence stock is well received, that is, if it contain characteristic vernacular or dialectal sentences that are correctly assigned areas. The attractiveness of the game comes from the widespread interest for one's own language and its variations. Experience shows that in all cultures there is a considerable interest in language issues and in reflecting on one's own language variations.

The agreement achieved among Borsa Parole players is adjustable: A player can revise at any time her sentence stock, especially refining the areas assigned to sentences and her estimates of how well these areas are likely to be recognised by other players. Thus, Borsa Parole taps in the self-regulation of a market for speakers themselves to generate linguistic estimates that steadily converges towards social agreements.

A salient feature of Borsa Parole is its self-sufficiency: Except at the game's start, no sentences have to be provided by the linguists pursuing the study. Both, the linguistic material and its interpretation is provided by the players themselves. This has two advantages. First, its makes the game self-sufficient and therefore scalable. There is no risk that a success of the game results in its laking of data and consequently in its loosing the interest of its players. Second, it ensures an unbiased data collection.

The contributions of this article are as follows:

- A novel GWAP based on a market providing an output-agreement strategy with adjustable agreement.

- A reflection on making the game self-sufficient and therefore scalable.

- A reflection on how to exploit GWAPs for linguistic field research.

This article is structured as follows. After this introduction, the linguistic motivation of this research is briefly presented. Then, related work is discussed. Then, the game is explained in more detail. Finally, perspectives for further research are mentioned in a conclusion.

Motivation: Field Research on the Divergence of the Italian Vernaculars and Dialects

The Italian language spoken today in all areas, cities and countryside alike, and within all social groups is currently experiencing a divergence. This makes the Italian language different from most other European languages and especially worth of linguistic field research.

During the restructuring and standardisation process which the Italian language experienced more recently than most other European languages only in the late 19th century, a common language emerged out of several rather disparate dialects. However, instead of being perceived as languages for less educated people, the Italian dialects have remained in today's spoken and written language across all social groups. A witness of this is strength of the Italian dialects is their presence on Wikipedia.

Vernaculars, that is unstandardised language varieties, too, are commonly used in Italy up till today. Vernaculars differ from (standardised) dialects and from each other in vocabulary,

grammar, and/or pronunciation. Standard Italian does not have a widely accepted pronunciation in contrast to English which gains it among others through the strict adherence of the BBC to a standard pronunciation [14]. Furthermore and importantly, the vernaculars spoken in large Italian cities currently evolve. Especially, new vernaculars emerge, disconnecting metropolises from one another [10].

Linguistic studies, especially on Italian vernaculars and dialects, require to gather multi-dimensional data on speakers', especially their social background, their geographical areas, and their education, the situation in which the speech takes place, like formal or informal, the time at which it takes place, and also whether thew language used is written or spoken [10]. So far, there is not much data data of this kind available for Italian vernacular and dialects.

The focus of the field research for which Borsa Parole has been developed is on vernacular and dialect sentences associated to geographical areas and speakers' genders, ages, and education level. In a first stage, only written language data is collected. Data on spoken languages will be gathered at a latter stage of the study.

Related Work

Von Ahn and Dabbish have called GWAPs games that use human computation to solving a computational problem [20]. This article proposed the ESP game in which the same image is shown to two randomly paired players who are rewarded if they suggest the same label for that image. Since the only resource shared by the two players is the image, the players tend to enter descriptions that are likely to be given also by their counterparty player. The image descriptions collected with the ESP Game have been used to improve image search.

Several other GWAPs have been designed that solve different goals, among others games for protein-folding [3] and for eliciting user preferences [5].

The verification strategies for asserting that a player's input can be accepted as a correct solution are limited if no algorithm can verify their correctness. The article [21] and the book [12] investigate possible verification strategies and suggests the following game-structure templates for two-player GWAPs: output-agreement games, inversion-problem games and input-agreement games. Most GWAPs that have been proposed so far fall into on one of these principles. More recently, additional verification strategies have been developed such as the complementary-agreement strategy of the game Polarity [11].

The GWAP Borsa Parole described in the present article differs from standard output-agreement games as characterised by von Ahn and Dabbish in [21] inasmuch that Borsa Parole players can also express an extent to which an agreement is to hold and also because this extent can be adjusted. We therefore call Borsa Parole an "output-agreement game with adjustable agreement".

Jain and Parkes have explored in [8] how incentives change the equilibrium of a game and therefore the outcomes it yields. Ho and Chen abstract in [7] from a specific game and propose formal models for verification. These two articles have influenced our design of Borsa Parole.

Munro et al. present in [16] linguistic projects exploiting Human Computation, specifically, Amazon Mechanical Turk, where users are paid for completing tasks. An important conclusion of this article is that the linguistic quality achieved using Human Computation is comparable to that of controlled laboratory studies. Further articles report on using GWAPs for gathering

corpora annotations [1, 6], transcriptions [17], and emotions, intentions, and attitudes [18].

Duolingo[1] is a platform offering its users support in learning languages which collects material for automating translation.

Borsa Parole, being a market, is related to prediction markets where prices are used to estimate the probability of future events [23]. Borsa Parole is similar to prediction markets inasmuch that the same self-regulation is used for generating a social agreement. Borsa Parole differs from prediction market inasmuch that, instead of predicting a future event, it serves to the collective assessment of properties of a language by the speakers of this language. Note that some researchers have expressed the view that direct estimates might be more precise than those generated on a prediction markets [15].

To the best the author's knowledge, no other GWAP than Borsa Parole have been proposed so far that relies on a market.

Gameplay

A player can play in two different and complementary manners that can be rather freely intertwined during the course of a play. A player can either enter new sentences and speculate on their properties (where and by whom they are spoken), or by speculate on sentences that have already been suggested by others or by oneself. Thus, the game really is a language stock exchange where a player can both sell, that is, submit sentences, and speculate on already suggested sentences.

In speculating on sentences already suggested, each session consists of five successive rounds during each of which a sentence is the object of focus. So far, the game provides written sentences but an extension with spoken sentences is foreseen.[2]

A sentence is presented to a player as shown by Figure 1 which can first specify in which area she thinks the sentence is used. Then, she can

- speculate on the proportion of players who, in her opinion, are likely to assign the same area to this sentence,

- highlight the words from the sentence that motivated her choice, and

- speculate on on the gender, age, and level of education of speakers of that sentence.

After five rounds are completed, a summary of the five sentences presented during the rounds is displayed to the player who can now inspect, and possibly re-assess her speculations.

In the following, some steps are discussed in more details.

In the first phase of a round, a player has to assign an area to the displayed sentence. This is done either by stepwise selecting an area on a map which, at each choice, focuses on the selected area and reduces its scale, or by entering the (beginning of the) name of an area into the text field. Instead of answering, a player can jump to the next sentence.

The player can also select a question mark next to the sentence in case it is not fully understandable to her. A translation in standard Italian is then given. The purpose of this translation

---

[1]http://www.duolingo.com

[2]The foreseen extension of Borsa Parole with spoken sentences does not require any change to the game logic.

Figure 1: Borsa Parole during the choice of an area for the displayed sentence (in English "Try to take this exam, whatever may happen!"). The player is in round ("Turno") 1 of 5 and achieved 0 points ("Punti") so far. "Viene da" means "Comes from", "Sud" means "South" (the current area selected), the buttons are "OK" and "Back", and the request in the red box on the left hand side means "For a more precise choice click on the map again, or confirm." The button above the map means "I don't know. Skip".

is to avoid a player's frustration and to increase the attractiveness of the game by making it a place where to discover vernacular and dialectal sentences. The software keeps track for later linguistic analysis of every asking for a translation as well as of every other actions performed by a player.

After assigning an area to a sentence, a player can speculate on the proportion of players she thinks are likely to assign that sentence to the same area as she does. Such a speculations are added to a player's list of speculations. From her speculations list, a player can track her performances and adjust the values and data she speculates with. As people guess, the values and data change and as a consequence, each speculation is valued differently over time. Points are granted for good matches with the actual values and data as explained below in Section "Game's Scoring".

The changing valuations are an incentive for players to review and adjust their speculations and to come back to the game. Indeed, their speculations give them rewards (points) even when they are not active on the platform.

A player is prompted to highlight the words from the phrase displayed that influenced her choice and further on to give clues on the gender, age, and level of education a speaker of the phrase is, in her opinion, likely to have. Note that this step can be skipped, so as the game not to be too demanding on the players. As with other speculations, a player gets feedback on how her estimations compare to the values observed. If a player's speculation is good enough, points are awarded to the player.

Tapping in a Market's Reflexivity

A player's speculation on a sentence consists of a five-tuple consisting of:

A Phrase
>   Which can be a complete sentence or an expression.

A Player
>   The player who speculates.

An Area
>   The geographic region, political region, province, or municipality where the player speculates that the phrase is spoken.

Speakers' Characteristics
>   Among other gender, age, level of education of the people that, in the player's opinion, are likely to use this phrase.

Recognition
>   The estimated proportion of players the player thinks are likely to recognise the phrase like she does.

The motivation for speculating on the market is that players are on the one hand animated to provide with data they perceive as correct, and on the other hand awarded for estimating how

many players agree with them. Doing so, they assess how much of an expert they consider themselves to be for this phrase and/or by setting the percentage to a high (common knowledge in Italy) or low (regional knowledge, uncommon, thus expert) value.

If a player is rather confident of an area for a phrase and thinks that other players are likely to confirm his assessment, she will choose a high recognition value. If a player is confident in her choice but believes that a phrase is unlikely to be often recognised, she will choose a low recognition value. If a player is unsure about her choice, she will either not speculate on this phrase or speculate with a low recognition value.

Unsure players are furthermore identified by their asking for translations in Standard Italian of the phrase displayed.

By allowing players to modify their speculations, the market regulates itself. A player can learn from other players and re-adjust her speculation so as to be more successful. In other words, the market incites players to express not what they might believe in but instead what they believe the community believes. This reflexivity in the behaviour of participants in a market has been called "beauty contest" effect by Lord Keynes [9] in his celebrated criticism of speculative markets. For our purposes, the beauty contest effect is no deficiency but instead an advantage. Indeed, all we wish to track is how a community assesses phrases. As of language, the believes of individuals are irrelevant.

Game's Self-Sufficiency

We call a GWAP self-sufficient if its generates the data its needs for its proper functioning. If a GWAP is not self-sufficient, then it must be regularly provided with new data. Especially if the GWAP becomes successful and attract always more players, a lack of self-sufficiency might be a serious problems and might threaten the game's scalability [19]. Indeed, providing a game with the additional data its increasing player's audience requires will soon require more human work the game operators can provide with resulting eventually in the game frustrating and therefore loosing its audience.

The data needed by a game can be provided in a self-sufficient way either by having access to an unbounded amount of data which is gradually offered by the players like for example the abundant amount of images available on the Web the ESP game [20] is based upon or by letting players themselves create the data other players process.

If the data needed by a game have to be added by some human work, then the game structure has to be adapted. The operators of Foldit have chosen to present only a few puzzles at a time, that is, not to scale the game's data with the number of its players [2].

Borsa Parole has been designed so as to be self-sufficient not only so as to avoid additional human work for gathering the data needed by the game but also for avoiding biases in the data collection. Indeed, vernaculars ad dialects that, so far, are insufficiently known, cannot easily be collected in an unbiased manner by the game's operators.

Borsa Parole has a cold start problem, though: Seed data are needed for the first players to have data to play with. We retain a standard bootstrapping approach for solving this problem which consists in gathering seed data by¡ –admittedly biased human work [22].
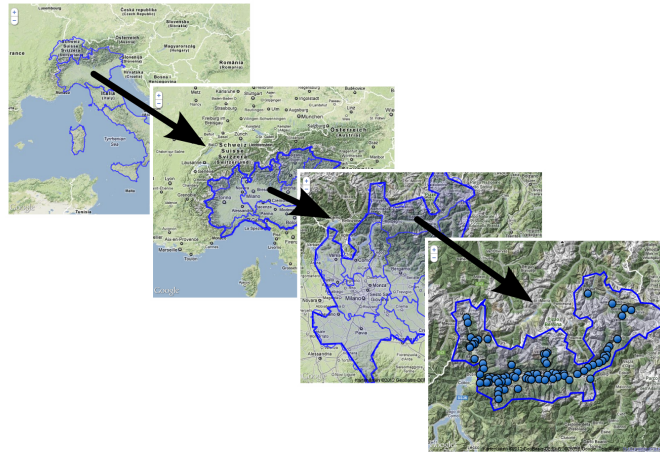
Figure 2: Stepwise focusing on smaller areas: broad geographic regions, political regions, provinces, and municipalities.

Game's Scoring

The scoring system is another integral part of a GWAP to achieve the intended goals. It provides players with feedback on their actions. A game's scoring should reward players for actions that are beneficial to the GWAP's objectives but not for actions that do not contribute to that goals. Furthermore, rewards should motivate the player to keep on playing.

To achieve this, most scoring of Borsa Parole is based on a player's output agreement with the rest of the community of players. For all speculations, the best concordance with values observed from the community provides the highest reward. The maximum score for a speculation is linked to the number of similar speculations: The more players speculate, the higher the maximum score.

For scoring area assignments, the geographic vicinity is important. Therefore, we also score nearby areas, however less than exact matches. The chosen areas can differ in the scale of the area (broad geographic region, political region, province, and municipality, see Figure 2) as well as the geographic vicinity within one scale (e.g., neighbouring provinces). If a player assigns an area which is adjacent to the area assigned by most other players, he still receives 80% of the points he could have got for an exact match. The same is true if he chooses Italian's capital city Rome whereas the majority assigned the province surrounding Rome (i.e., a more precise assignment). For choosing the province Rome if the majority of the other players assigned the city Rome he receives 50% of the points because this information is less detailed but still useful.

Conclusion and Perspectives

This article has described a novel approach to linguistic field research using a GWAP called Borsa Parole of a novel kind based on a market and speculation paradigm. Borsa Parole relies on

the beauty contest effect of speculative market for fostering and tracking a community agreement of properties of vernaculars and dialects.

The use of Human Computation, including GWAPs, in linguistic has just begun. Arguably, Human Computation is very promising as a new, cheaper and less biased form linguistic field study than traditional linguistic field study.

Acknowledgements

References

[1] Jon Chamberlain, Massimo Poesio, and Udo Kruschwitz. Phrase Detectives: A web-based collaborative annotation game. In *Proceedings of the 4th International Conference on Semantic Systems (I-SEMANTICS)*, 2008.

[2] Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović, and Foldit Players. Predicting protein structures with a multiplayer online game. *Nature*, 466(7307):756–60, 2010.

[3] Seth Cooper, Adrien Treuille, Janos Barbero, Andrew Leaver Fay, Kathleen Tuite, Firas Khatib, Alex Cho Snyder, Michael Beenen, David Salesin, David Baker, and Zoran Popović. The challenge of designing scientific discovery games. In *Proceedings of the 5th International Conference on the Foundation of Digital Games (FDG)*, pages 40–47, 2010.

[4] Kathryn A. Davis. Qualitative theory and methods in applied linguistics research. *TESOL Quarterly*, 29(3):427–453, 1995.

[5] Severin Hacker and Luis von Ahn. Matchin: Eliciting user preferences with an online game. In *Proceedings of the 27th Annual CHI Conference on Human Factors in Computing Systems*, pages 1207–1216, 2009.

[6] Barbora Hladká, Jií Mírovský, and Pavel Schlesinger. Designing a language game for collecting coreference annotation. In *Proceedings of the 3rd Linguistic Annotation Workshop (ACL-IJCNLP)*, pages 52–55, 2009.

[7] Chien-Ju Ho and Kuan-Ta Chen. On formal models for social verification. In *Proceedings of the Human Computation Workshop*, pages 62–69, 2009.

[8] Shaili Jain and David Parkes. A game-theoretic analysis of games with a purpose. In *Proceedings of the 4th International Workshop on Internet and Network Economics (WINE)*, pages 342–350, 2008.

[9] John Maynard Keynes. *The General Theory of Employment, Interest, and Money*. Macmillan Cambridge University Press, 1936.

[10] Thomas Krefeld. Italienische Varietätenlinguistik. *Italienisch. Zeitschrift für italienische Sprache und Literatur*, 63:56–62, 2010. In German.

[11] Edith Law, Burr Settles, Aaron Snook, Harshit Surana, Luis von Ahn, and Tom Mitchell. Human computation for attribute and attribute value acquisition. In *Proceedings of the First Workshop on Fine-Grained Visual Categorization (FGVC)*, 2011.

[12] Edith Law and Luis von Ahn. Human Computation. In Ronald J. Brachman, William W. Cohen, and Thomas Dietterich, editors, *Synthesis Lectures on Artificial Intelligence and Machine Learning*, pages 1–121. Morgan & Claypool Publishers, June 2011.

[13] Anne Lazaraton. Qualitative research in applied linguistics: A progress report. *TESOL Quarterly*, 29(3):455–472, 1995.

[14] Anna Laura Lepschy and Guilio Lepschy. *The Italian Language Today*. Routledge, 2nd edition, 1988.

[15] Charles F. Manski. Interpreting the predictions of prediction markets. Technical report, National Bureau of Economic Research, 2004.

[16] Robert Munro, Steven Bethard, Victor Kuperman, Vicky Tzuyin Lai, Robin Melnick, Christopher Potts, Tyler Schnoebelen, and Harry Tily. Crowdsourcing and language studies: The new generation of linguistic data. In *Proceedings of the NAACL-HLT Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 122–130, 2010.

[17] Tim Paek, Yun-cheng Ju, and Christopher Meek. People Watcher: A game for eliciting human-transcribed data for automated directory assistance. In *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2007.

[18] Lisa Pearl and Mark Steyvers. Identifying emotions, intentions, and attitudes in text using a game with a purpose. In *Proceedings of NAACL-HLT Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Los Angeles, CA, USA, 2010.

[19] Katharina Siorpaes and Martin Hepp. Games with a Purpose for the Semantic Web. *IEEE Intelligent Systems*, 23(3):50–60, 2008.

[20] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 319–326, 2004.

[21] Luis von Ahn and Laura Dabbish. Designing games with a purpose. *Communications of the ACM*, 51(8):58–67, August 2008.

[22] Jakub Šimko and Mária Bieliková. Games with a purpose: User generated valid metadata for personal archives. In *Proceedings of the 6th International Workshop on Semantic Media Adaptation and Personalization*, pages 45–50, December 2011.

[23] Justin Wolfers and Eric Zitzewitz. Prediction markets. *Journal of Economic Perspectives*, 18(2):107–126, 2004.