

# Dealing with DEAL: A dialogue system for conversation training

**Anna Hjalmarsson**  
Centre for Speech Technology  
KTH, Stockholm,  
Sweden  
annah@speech.kth.se

**Preben Wik**  
Centre for Speech Technology  
KTH, Stockholm,  
Sweden  
preben@speech.kth.se

**Jenny Brusk**  
Department of Game  
Design, Narrative and  
Time-Based Media  
Gotland University,  
Sweden  
jenny.brusk@hgo.se

## Abstract

We present DEAL, a spoken dialogue system for conversation training under development at KTH. DEAL is a game with a spoken language interface designed for second language learners. The system is intended as a multidisciplinary research platform where challenges and potential benefits of combining elements from computer games, dialogue systems and language learning can be explored.

## 1 Introduction

There is a growing trend among educational researchers to look at games and game design in order to make education more appealing and effective. A new and challenging domain for spoken dialogue systems is *serious games*, i.e., applications of interactive technology that have purposes other than solely to entertain, including training, advertising, simulation, or education (Iuppa & Borst, 2007). If successful, serious games will engage users motivated by a willingness to be entertained and/or educated. Encouraged by such motivations users will be prepared to talk to dialogue systems because it is fun, repeatedly and for long periods without the need for predefined tasks. This is a tempting scenario.

We present DEAL, a spoken dialogue system for second language learners of Swedish under development at KTH. DEAL is intended as a multidisciplinary research platform where challenges and potential benefits of combining elements from computer games, dialogue systems and language learning can be explored. From a

dialogue research point of view a serious game approach contributes with several novel and interesting objectives and challenges. These include how to design dialogues which are fun and natural using a language which suits the vocabulary and language complexity of language learning students on various levels. Since efficiency and task completion are no longer the main objectives, dialogue systems in a serious game context do not have to be predictable, rational or even co-operative. Instead, we need to consider how to build systems which are fun, educational and addictive to talk to.

### 1.1 Acquiring conversational skills

Language learning can be modelled as a series of developmental steps going from declarative to procedural knowledge. First, an item is noticed in a meaningful contrastive situation, then it occurs repeatedly in meaningful input and is practised in communication until it is internalised, and finally automatised (Ellis, 2006). To automatise these processes when learning a second language we need a meaningful situation where conversational skills can be practised repeatedly. Because of its complexity, learning a language requires substantial effort and the motivation varies both over time and between individuals. To practise conversational skills while playing a game may increase any existing motivation to learn if there is one, and creates a motive to learn if there isn't. Our objective is similar to the Nice project (Gustafson et al., 2004), in that we wish to create a game where spoken dialogue is not just an add-on, but is used as the primary means for game progression.

## 2 Motivation

The practical motivation of DEAL is to build an application where conversational skills can be practised in a fun and meaningful context. In short, DEAL is a game with a spoken language interface designed for second language learners. A similar approach is used in the tactical language training system (TLTS), a large-scale application that helps people acquire basic conversational skills in Levantine and Iraqi Arabic (Johnson et al., 2005). Our first choice of domain for this work is the trade domain. DEAL sets the scene of a flea market where a talking animated agent is the owner of a shop where used objects are sold. The domain was chosen for several reasons:

- A trading situation is a fairly restricted and universally well-known domain. It is something everyone is conceptually familiar with, regardless of cultural and linguistic background.
- A trading situation is from a language learning point of view a very useful domain to master in the new language
- The objects sold at a flea market can be a diverse set of items which can be tailored to suit the vocabulary mastered by a language learning student.
- A flea market is a place where it is acceptable to negotiate about the price. Negotiation is a complex process which includes both rational and emotional non-rational elements. This opens up for interesting and complex dialogue.

These characteristics combined gives us an application where users can engage in a dialogue situated in a well-known context but which also includes elements of surprise and challenge (i.e., getting a good price).

### 2.1 Ville

DEAL is developed as a free-standing part of Ville, a framework for language learning developed at KTH (Engwall et al., 2004). Ville is a virtual language tutor helping students to improve their listening and pronunciation skills in a new language. Ville detects and gives feedback on pronunciation errors, and has challenging exercises that are used in order to teach new vocabulary, or

to raise the students' awareness of particular perceptual differences between their first and second language. Ville has exercises on phone, syllable, word and sentence level.

DEAL adds the possibility to give conversation training. Whereas Ville is a language tutor who provides the user with feedback on performance, the agent in DEAL does not comment on your performance but acts as your conversation partner in a role-playing fashion. Using DEAL as an integrated part of Ville, the system has knowledge about particular students' acquired vocabulary. This information can be used to tailor the language in DEAL as well as the items being sold.

## 3 Implementation

DEAL is implemented using components from the Higgins project (Skantze, 2005), an off-the-shelf ASR system, a dialogue manager developed for DEAL purposes and a GUI with an embodied conversational agent (ECA).

### 3.1 User interface

Our ECA (embodied conversational agent) is developed at KTH (Beskow, 2003), and can use either synthetic or natural, pre-recorded speech. The head is capable of producing lip-synchronized speech as well as extra linguistic signs such as frowning, nodding, and eyebrow movements. Language is multimodal, and in second language learning, visual signals are an important source of information.

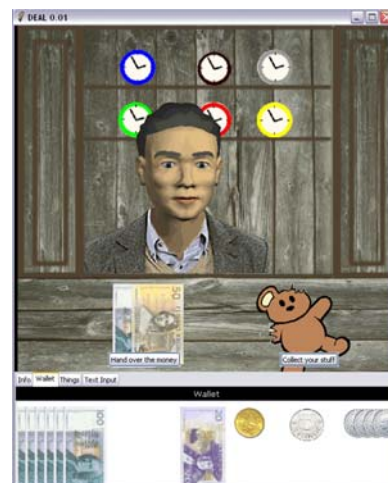


Figure 1: DEAL user interface

Higgins includes modules for semantic interpretation and analysis. Pickering, a modified chart parser, supports continuous and incremental input from a probabilistic speech recognizer. Speech is unpredictable and chunking a string of words into utterances is difficult since pauses and hesitations will likely be incorrectly interpreted as end of utterance markers. This will be even more evident for second language learners whose conversational skills are not yet automatized and whose language contains disfluencies such as hesitations and false starts. Pickering uses context free grammars (CFG) and builds deep semantic tree structures. Grammar rules are automatically relaxed to handle unexpected, ungrammatical and misrecognized input robustly. The discourse modeler, Galatea, interprets utterances in context and keeps a list of the communicative acts (CA) in chronological order. Galatea resolves ellipses, anaphora and has a representation of grounding status which includes information about who added a concept, in which turn a concept was introduced and the concept's ASR confidence score.

#### 4 The DEAL domain

Game designers focus on finding ways to keep players engaged and motivated throughout a game. Nonetheless, dialogues in today's games have a strict way of affecting the continuance of the game. The interaction is typically based on complex tree structures, where one action leads to a set of new choices. Choosing one line or topic has an immediate result and the dialogue traverses a finite branching tree structure. With these types of dialogues it is fairly trivial how to get the desired result, making it less interesting to engage in the interaction. We strive towards an interaction with a less predictable result. Façade is an interactive drama project that introduces a drama manager to make the outcome of a dialogue less predictable (Mateas & Stern, 2003). In Façade the story is divided into beats, an atomic unit of drama, where beats and transitions between beats can unfold in various ways depending on what type of input is provided by the user.

##### 4.1 Dealing with DEAL

DEAL has two actors, one ECA and one human language student. The student is given a mission to buy items at a flea market getting the best possible

price from the odd looking shop-keeper. The shop-keeper can talk about objects and their properties and negotiate about the price of the objects. The most challenging part in DEAL, both from a "buyer" (user) point of view and when designing the conversational agent, is negotiating about the price of objects. At first, dealing about price can seem like a fairly rational and straight forward procedure. However, negotiating is a complex multidisciplinary area of research which touches fields such as psychology, economics and political science. Negotiating about a price in a face to face situation involves a number of various parameters which are often affected by non-rational and emotional aspects. Second hand items may have rich interesting characteristics which makes them interesting to talk about. For example the items can be defective, have a personal history or an affection value to the shop-keeper, all of which may have an impact on the negotiation process.

The dialogue can unfold in different ways depending on what the user says (see Figure 2). Negotiation is implemented using a fairly straight forward algorithm and a few heuristics. To introduce elements of gameplay we have integrated a parameter which represents the agent's "willingness" to reduce the price of an item. The willingness parameter is the percentage share of the seller's original price that the ECA is willing to accept, after negotiating, as price for a particular item. The parameter has an initial value which may be affected depending on how the dialogue proceeds. To affect the outcome of the interaction, the player may try to influence the willingness of the shop-keeper to reduce the price.

- U1: I'm interested in buying a toy.  
 S1: Oh, let me see. Here is a doll.  
 (a doll is displayed)  
 U2: Do you have a teddy-bear?  
 S2: Oh, yeah. Here is a teddy-bear.  
 (a teddy-bear is displayed, see Figure 1)  
 U3: How much is it?  
 S3: You can have it for 180 SEK  
 U4: I give you 1 SEK (*willingness decrease*)  
 S4: No way! That is less than what I paid for it.  
 U5: Ok how about 100?  
 S5: Can't you see how nice it is?  
 U6: But one ear is missing. (*willingness increase*)  
 S6: Ok, how about 150?  
 U7: 130?  
 S7: Ok, it is a deal!

Figure 2: Dialogue example from DEAL

The outcome of the game is affected by what the user says. For example in utterance U4 the seller is offended by the user's low bid and his willingness to give the user a good price is reduced. However, when the user points out a flaw of the object (the GUI displays a teddy-bear with one ear, see Figure 1) the seller feels obligated to give the user a better price, i.e., his willingness increases.

#### 4.2 Dialogue characteristics in DEAL

Humans who engage in a dialogue tend to coordinate their linguistic behaviour (Pickering & Garrod, 2006), sometimes referred to as entrainment. Research on linguistic entrainment in human-machine interaction has shown that users of spoken dialogue systems also adopt the system's way of speaking (see for example Brennan, 1996). Moreover, research and literature on second language acquisition (SLA) is diverse, with no single theory or model seen as the most appropriate. However, there seem to be a consensus about the value of conversational interactions. The more you talk the better it is.

Consequently, from a second language learning perspective, the language used in DEAL will be crucial. It is important that the agent behaves human-like in a way which motivates the users to talk a lot and not only in short command-like utterances. The goal is not to create a conversational agent which behaves human-like in every sense but which is human enough to make the users suspend their disbeliefs, i.e. make them act as if they were talking to another human being (Cassell, in press). This does not necessarily mean that the agent needs to be cooperative or polite. The seller can actually be rude and try to avoid the users' requests as long as this is done in a way that does not destroy the users' willingness to accept the ECA in DEAL as a character with human-like conversational capabilities.

#### 5 Concluding remarks

Whether DEAL is a fun game or not is yet to be investigated. So far, the scenario, rules and possible actions in DEAL are fairly limited. Much can be added to the system in the long run, but this far our main motivation has been to introduce simple examples of social interaction that affect game progression.

#### 6 Acknowledgements

This research was carried out at Centre for Speech Technology, KTH. The research is also supported by the Graduate School for Language Technology (GSLT). Many thanks to Rolf Carlson, Gabriel Skantze, Joakim Gustafson, Jens Edlund.

#### References

- Beskow, J. (2003). *Talking Heads - Models and Applications for Multimodal Speech Synthesis*. Doctoral dissertation, KTH.
- Brennan, S. E. (1996). Lexical entrainment in spontaneous dialog. In *Proceedings of ISSD*.
- Cassell, J. (in press). Body Language: Lessons from the Near-Human. In J. Riskin (ed.) *The Sistine Gap: History and Philosophy of Artificial Intelligence*. Chicago: University of Chicago Press.
- Ellis, N. (2006). Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Saliency, Interference, Overshadowing, Blocking, and Perceptual Learning. *Applied Linguistics*, 27, 164-194.
- Engwall, O., Wik, P., Beskow, J., & Granström, G. (2004). Design strategies for a virtual language tutor. In Kim, S. H. & Young, D. H. (Eds.), *Proc ICSLP 2004* (pp. 1693-1696). Jeju Island, Korea.
- Gustafson, J., Bell, L., Boye, J., Lindström, A., & Wirén, M. (2004). The NICE Fairy-tale Game System. In *Proceedings of SIGdial*. Boston.
- Iuppa, N., & Borst, T. (2007). *Story and simulations for serious games : tales from the trenches*. Focal Press.
- Johnson, W., Vilhjalmsson, H., & Marsella, S. (2005). Serious games for language learning: How much game, how much AI?. In *12: th International Conference on Artificial Intelligence in Education*. Amsterdam.
- Mateas, M., & Stern, A. (2003). Façade: An experiment in building a fully-realized interactive drama. In *Game Developer's Conference: Game Design Track*. San Jose, California, US.
- Pickering, M., & Garrod, S. (2006). Alignment as the Basis for Successful Communication. *Research on Language and Computation*, 4, 203-228.
- Skantze, G. (2005). Galatea: a discourse modeller supporting concept-level error handling in spoken dialogue systems. In *Proceedings of SigDial* (pp. 178-189). Lisbon, Portugal.