

TALERUM - learning Danish by doing Danish

Peter Juel Henriksen

Danish Center for Applied Speech Technology

pjh.msc@cbs.dk

Abstract

Talerum is an interactive L2 learning environment, here presented publicly for the first time. Talerum is designed for use in Danish language classes throughout the vast and sparsely populated West-Nordic area. In a town-like setting (resembling Copenhagen) the pupil navigates among shops (foods, clothes, tools), a cafe, a school and, importantly, a home base where the game opens and the pupil is welcomed as an exchange student. Through elaborate dialogues with the game characters, the pupil discovers and pursues a *secret mission*. Talerum thus uses elements of game-logic, informal dialogue, and relatively deep semantic analysis to get and keep the attention of the language student of generation IT.

1 Background

Talerum is the last corner stone in the suite of language training tools developed and user-tested by the working group FRASAR. The entire suite is hosted by Iceland University at

<https://taleboblen.hi.is>

Established in 2012, the Frasar group (Nordic computational linguists and didacticians) reaches out to language teachers and pupils in the vast West-Nordic area (Greenland, Iceland, the Faroese Islands) where Danish is taught as a second/third language. The Frasar tools support the productive aspects of language learning, especially informal Danish as used in everyday situations (smalltalk, friends' conversations, greetings, shopping, ...). The tools are browser-based and combine gaming elements, example dialogues (with recordings by native Danes), spoken language exercises (acoustic analysis of pupils' pronunciation), songs, jokes, puzzles and interactivity to stimulate language awareness and prepare the pupil to meet the Danes. See Lihong (2013) and Chik (2014) for perspectives on gaming as a strategy for L2 teaching wrt. individual and social learning.

2 Talerum

At NODALIDA-17 we will demonstrate Talerum, perhaps the most ambitious Frasar tool. Talerum offers a free-style exploration of a virtual suburb with shops, a cafe, a school, and a home base. Pupils enter the game with little or no formal preparation. The intended, but untold, plan is for the pupil to first explore the game universe aimlessly, then uncover a hidden mission, and finally work systematically on solving her task to earn a reward of points and praise. Each step requires (and rewards) concise and relevant Danish-language productions in numerous dialogues with the game characters. While most of the Frasar tools aim at language correctness, Talerum allows the pupil to *do* things with language: present themselves to a stranger, obtain information through dialogue, negotiate with a shop attendant, and even smalltalk. Contrary to most CALL tools (Johnson 2005, Godwin-Jones 2014, Chun 2016) Talerum combines shallow syntax analysis (not penalizing syntax errors) with fairly deep semantic evaluation. See Berns (2012) and Chik (2014) for relevant discussion. Also we admit to being inspired by the classical Larry Laffer game.

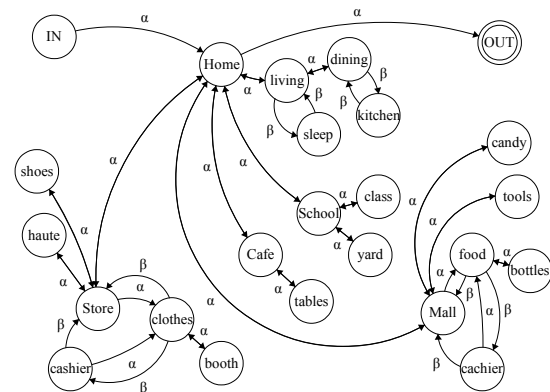


Figure 1. Talerum scenes and gates. All sessions begin and end at the family home (state "Home"). α -gates are always open, β -gates depend on the pupil's current game status.

2.1 Interaction modalities

Pupil clicks a gate button. **System** responds by changing the scene (e.g. "Prøverum" in fig.2).

Pupil initiates or continues a dialogue with one of the Talerum characters, e.g. an opener "Dav, jeg hedder Thor", *hello I'm Thor*, or a specific query "har du et par sorte cowboybukser i small?", *do you have pair of black jeans size S?* **System** returns a Danish phrase in response.

Pupil clicks on a former phrase. **System** responds by playing back the phrase in speech synthesis (prosody of Danish vernacular: high speech rate, falling contour, phonetic reductions).

Pupil clicks a help button. **System** responds by supplying context-relevant information, either language assistance (specific Danish terms or phrases, with translations in the pupil's L1) or game-related advice (usually 'indirect', e.g. "I Danmark betaler næsten alle med dankort", *Most Danes pay their bills with the dankort*, a hint to the user to search for and collect a paycard).

2.2 Environments, scenes and gates

Talerum has five main environments, viz. a home base, a school, a café, a department store, and a food court (see fig. 1, Home, School, Cafe, Store, Mall, respectively). The environments serve as hubs, each connecting a sub-system of scenes (lowercased in the graph). Some scenes are static, others dynamic. Static scenes only allow entry and exit (they serve to widen the navigation space and add a little suburb realism, liked by our child testers). Passage through static scenes does not contribute to the game mission and is neither rewarded nor punished wrt. performance evaluation. Dynamic scenes, in contrast, involve dialogue potentially promoting the mission.

The most elaborate interaction takes place in the home environment (living room, sleeping room, kitchen), the food market, and the clothing shop SuperChik (fig. 2). The latter is a second-hand store for fashion items (becoming ever more popular in Copenhagen), the pupil taking the role as a customer enquiring about particular items specifying her wishes and negotiating quality and price. She may even change her mind or ask broader questions not relating to clothing. The system's replies are generated by a range of strategies including canned lines ("Hvad skulle det være?"), phrase

templates ("Vi fører desværre ikke X"), and genuinely compositional constructions (introduced below). On the user side, Talerum's spelling control is strict, syntax control relaxed.

As illustrated in fig.2, the three most recent turns remain visible. The line "Dig før: en blå kjole" (*You before: a blue dress*) quotes the pupil (and also links to the TTS playback). The next line is the attendant's turn (here describing a particular dress). The third line "Din tur: ja tak, den tager jeg" (*Your turn: OK, I'll take it*) has just been entered by the pupil, but not sent yet. Pressing "Sig det!" at this point will have the effect of a classical speech act. The attendant will acknowledge the sale transferring the commodity to the pupil's basket, while the gate panel (right side) will change accordingly replacing gate "Megazin" (building exit) by a new gate "Kassen" (*Cashier*, fig.3).



Figure 2. SuperChik, a second-hand clothing shop in the Department Store environment.



Figure 3. The cashier's desk. Access to this scene requires an item in the pupil's shopping basket. Exit from this scene to the Department Store's main hall requires a successful payment.

2.3 Talerum dialogue - an example

Pupil:- Dav, jeg vil gerne se på et par støvler.
[hello, I'd like to see some boots]

Clerk:- Vi fører desværre ikke fodtøj. Var der andet, du kunne tænke dig?
[I'm sorry, but we don't have footwear; anything else?]

Pupil:- Så vil jeg gerne have nogle bukser.
[then I'd like to look at some trousers]

Clerk:- Ja tak. Vi har mange forskellige slags, var der noget bestemt du kunne tænke dig?
[we have several kinds, anything in particular you'd like?]

Pupil:- Gerne hvide, og helst af ren bomuld
[white is best, and of pure cotton if possible]

Clerk:- Hvad med den her, en lækker hvid buks i bomuld til 149 kroner. Noget for dig?
[what about this, a nice pair of white jeans size M, at 149]

Pupil:- kan jeg prøve den på?
[can I try them on?]

Clerk:- ja, prøverummet er lige her ved siden af
[certainly, the booth is right next to us]

Pupil:- [kommer tilbage] De er lidt for store
[they're somewhat large]

Clerk:- Her er en skøn bomuldsbukse i small, nok mest til dagligbrug, til 89 kroner. Var det noget?
[what about this lovely cotton, size S, at 89, you like it?]

Pupil:- ja tak, den tager jeg.
[yes please, I'll take them]

Clerk:- mange tak. Du kan betale i Kassen.
[thank you very much, please proceed to the cashier]

2.4 Semantic evaluation

Dialogues in the Home environment follow a predefined sequence of stages where each production may or may not allow a step forward.

In contrast, dialogues in the shops are user-driven, and responses are generated with inference-based methods. The derived discourse-information is accumulated (until the pupil leaves the scene) enabling the use of pronominal anaphor, stepwise clarification of wishes and offers, re-negotiation, and even intermediate off-topic enquiries.

The response generator takes as input (i) the pupil's latest production (normalized by filtering out function words etc.), (ii) a stack of propositions representing the preceding discourse (may contain contradicting information after re-negotiation, such as "size medium" and "larger than medium", giving priority to the most-recent), (iii) the set of all entities (goods), (iv) a no-go list of entities already rejected or sold, (v) a focus entity (only defined if a particular entity is being negotiated). The resulting type-logical

expression is mapped to a Danish phrase reflecting its type: a yes/no-question (e.g. "Kan du lide denne her?", *do you like this one?*), a material question ("Hvad kunne du ellers tænke dig?", *what else would you like?*), a statement ("Vi har ikke flere kjoler i den farve", *we have no more dresses in that color*), or a simple boolean (e.g. "Ja tak", *yes please*, or "Nej, desværre", *no we don't*).

The inference system is written in Prolog, the language-specific and www-related parts in Perl, Python and Linux-shell. Client-side code is html5 (including css and JavaScript). The overall architecture is module-based, so most functional modifications require local changes only, such as commodity types and properties, scenes&gates, dialogues, characters, and more. Portation of the entire Talerum to a new language locale only requires rewriting a single module (the phrase database used by the response generator) and replacing the lexicon file (word forms *cum* PoS). More technological details are to be published soon. Do ask us at NODALIDA.

2.5 The challenge

We have tested Talerum with children during the construction phase. Most like shopping around randomly and trying their hand with Danish dialogue for a while. However, after a while they loose interest, so we added a secret challenge. Through dialogue with your roomy (called *Emil* for boy users, *Ida* for girls, same age as the pupil) and family members you learn the details of your task, eventually sending you to town to obtain something (e.g. a present) making someone relieved or surprised. Completing the task will earn you a seasoned appreciation (depending on your time score and efficiency).

3 Concluding remarks

Talerum is currently being tested in folkeskoler and gymnasier in Kalaallit Nunaat, Ísland and Føroyar (pupils 11-14Y). The results (metadata and performance data) will be released in late 2017 and might be interesting for comparative studies. We already know from earlier work (Henrichsen 2015A&B, forthcom.) that Icelandic and Greenlandic pupil groups differ markedly wrt. proficiency, motivatedness, and linguistic challenges faced with the Danish language.

Talerum (ver. 2.0) will be released soon for classroom use. For practical info contact FRASAR's founder and chair Auður Hauksdóttir.

References

- Anke Berns, A Gonzalez-Pardo, D. Camacho. 2012. Game-like Language Learning in Virtual 3D Environments. 2012. *Journal of Computers and Education*. 60(1):210-220.
- Alice Chik. 2014. Digital Gaming and Language Learning: Autonomy and Community. *Journal of Language Learning and Technology*. 18(2): 85-100
- Dorothy Chun, R. Kern, B. Smith. 2016. Technology in Language Use, Language Teaching, and Language Learning. *The Modern Language Journal*. 100(S1):64-80.
- Robert Godwin-Jones. 2014. Games in Language Learning: Opportunities and Challenges. *Journal of Language Learning and Technology*. 18(2): 9-19.
- Peter Juel Henrichsen. 2015A. *Talebob – an Interactive Speech Trainer for Danish*. NODALIDA 2015.
- Peter Juel Henrichsen. 2015B. *Taking the Danish Speech Trainer from CALL to ICALL*. NLP4CALL (workshop at NODALIDA 2015).
- W. Lewis Johnson, Hannes Vilhjalmsson, and Stacy Marsella. 2005. *Serious Games for Language Learning: How Much Game, How Much AI?* C.-K. Looi et al. (Eds.) IOS Press.
- Cao Lihong. 2013. Development and Prospect of Online Language Learning, *Journal of Xiangnan University*. 13(03)