Conversational agents for language learning: state of the art and avenues for research on task-based agents

Serge Bibauw¹,²,³ • Thomas François² • Piet Desmet¹

¹ KU Leuven Kulak • iMinds • ITEC
² Université catholique de Louvain • CENTAL
³ Universidad Central del Ecuador
Dialogue systems for language learning

- What?
  - Dialogue systems
- Why?
  - What do we know?
- How?

DISCIPLINES AND TERMS
DEFINITION AND COMPONENTS
RATIONALE
EFFECTIVENESS STUDIES
TECHNOLOGICAL PROCESS
What?

DISCIPLINES AND TERMS

https://www.chatbots.org/gallery/
<table>
<thead>
<tr>
<th>Dialogue System?</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Chatbot</td>
</tr>
<tr>
<td>▪ Conversational agent</td>
</tr>
<tr>
<td><strong>Dialogue System</strong></td>
</tr>
<tr>
<td>▪ Spoken dialogue system</td>
</tr>
<tr>
<td>▪ Dialogue-based CALL</td>
</tr>
<tr>
<td>▪ Dialogue simulation</td>
</tr>
<tr>
<td>▪ Dialogue game</td>
</tr>
<tr>
<td>▪ Simulated dialogue</td>
</tr>
<tr>
<td>▪ Speech-interactive CALL</td>
</tr>
<tr>
<td>▪ Conversational language tutoring system</td>
</tr>
<tr>
<td>▪ Conversation simulator</td>
</tr>
<tr>
<td>▪ Communication-based language learning</td>
</tr>
<tr>
<td>▪ Chatterbot</td>
</tr>
<tr>
<td>▪ Conversational companion</td>
</tr>
<tr>
<td>▪ Conversational system</td>
</tr>
<tr>
<td>▪ Conversational ICALL system</td>
</tr>
<tr>
<td>▪ Dialogue-based conversation tutoring</td>
</tr>
<tr>
<td>▪ Dialogue program</td>
</tr>
<tr>
<td>▪ Interactive pedagogical drama</td>
</tr>
<tr>
<td>▪ Microworld interaction</td>
</tr>
<tr>
<td>▪ Virtual agent</td>
</tr>
<tr>
<td>▪ Virtual human</td>
</tr>
<tr>
<td>▪ Pedagogical agent</td>
</tr>
<tr>
<td>▪ Voice-interactive CALL...</td>
</tr>
</tbody>
</table>
Systematical literature study: Corpus collection

- Systematical search on Web of Science, Scopus and ProQuest [results: 604 / 494 / 1003 hits]
- Ancestry (citing) and forward (cited in) search
- Final corpus: 135 published and peer-reviewed papers from 1982 to 2015 (April)
<table>
<thead>
<tr>
<th>Year</th>
<th>Number of published papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>0</td>
</tr>
<tr>
<td>1990</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
</tr>
<tr>
<td>2010</td>
<td>9</td>
</tr>
</tbody>
</table>

**Type**

- Dissertation
- Book chapter
- Conference paper
- Journal article

Studies published on DSLL
Disciplinary approaches

**SPEECH/NLP**
Spoken dialogue systems
Embodied conversational agents
- [SPELL system]
- [DEAL]
- [Sprinter]
- [Let's Go]
- [SCILL]
- [IVELL]...
- 1970’s – 1999 – ...

**CALL**
Intelligent tutoring systems & ICALL
- [FAMILIA]
- [SPANLAP]
- [Die Sprachmaschine]
- [CALLE project]
- [Miniprof]...
- 1982 – 1999

**GAMES**
Games and virtual worlds
- [Alelo’s TLCT/Tactical Iraqi]
- [Edubba]
- [IDI Virtual Conversations]
- [Façade]
- [Restaurant Game]...
- ... – 1988 – ...

**AI/AMATEUR**
Chatbots
- [Dave ESL]
- [CSIEC]
- [Tutor Mike]
- [Eliza]
- [ALICE]
- [Jabberwacky]...
- 1966 – 2000 – ...

Focused on:
- Technological challenges of dialogue management (initially: mainly spoken)
- Interactivity and content creation
- Corrective feedback provision
- AI development in light of the Turing Test (mainly written)
Disciplinary approaches

Papers published on DSLL by field of study

<table>
<thead>
<tr>
<th>Field</th>
<th>Year</th>
<th>Number of published papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Games</td>
<td>1980</td>
<td>0</td>
</tr>
<tr>
<td>Educational technologies</td>
<td>1990</td>
<td>5</td>
</tr>
<tr>
<td>Computer-assisted language learning</td>
<td>2000</td>
<td>10</td>
</tr>
<tr>
<td>Natural language processing</td>
<td>2010</td>
<td>20</td>
</tr>
<tr>
<td>Speech technologies</td>
<td>2020</td>
<td>30</td>
</tr>
<tr>
<td>Computer sciences</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dialogue systems

Definition and Components
Different systems?

- **Spoken / Written / Multimodal**
  e.g. spoken dialogue systems vs. chatbots

- **Technologically simple / complex**
  e.g. chatbots vs. dialogue systems (Klüwer, 2011)

- **Focus on form / Focus on meaning**
  e.g. tutorial CALL vs. “communication-based CALL” (Murray, 1995)

**Dialogue activity**
- conversation, talk, information exchange, questions-answers, requests-grants, etc.

**with an automated agent**
- as interlocutor, as tutor, as character...
- virtual, embodied, robot, etc.
Dialogue activity

- conversation, talk, information exchange, questions-answers, requests-grants, etc.

with an automated agent

- as interlocutor, as tutor, as character...
- virtual, embodied, robot, etc.

Dialogue system for language learning (DSLL)
Practice English with Mike

Mike is very smart. You can ask him any questions. When you want to practice the conversations of the topic you are interested in, go to www.eslfast.com/robot/, type in "Let's practice topic ..." or "I'd like to practice topic ..." You must type the exact name of the topic. For example, Let's practice topic Weather (2). When you are acting as A and Robot Mike is B, you must type in the exact sentence structure of the conversation, including the punctuation. If you make a mistake, it is not controlled practice of the conversation anymore; you need to type "Let's practice topic ..." again to get the robot back on track.

YOU: Hi Mike
Mike: Hello!
YOU: What do you know?
Mike: I'd rather not tell you.
YOU: Can you tell something insightful for the people at CALICO?
Mike: Sure.
YOU: Go on.
Mike: Just a moment.
YOU: We're listening.
Mike: You are?
SPELL system (Morton, Gunson & Jack)

### Distinctive features

<table>
<thead>
<tr>
<th>Interlocutor</th>
<th>Interaction</th>
<th>Input</th>
<th>Interactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>Monologue</td>
<td>Constrained</td>
<td>Fixed</td>
</tr>
<tr>
<td>Human + Automated tutor</td>
<td>“Mono-initiative” dialogue</td>
<td>Partially constrained</td>
<td>Fixed with illusion of choice</td>
</tr>
<tr>
<td>Automated interlocutor</td>
<td>Dialogue</td>
<td>Unconstrained</td>
<td>Interactive</td>
</tr>
</tbody>
</table>

- **Interlocutor**
  - Human
  - Human + Automated tutor
  - Automated interlocutor

- **Interaction**
  - Monologue
  - “Mono-initiative” dialogue
  - Dialogue

- **Input**
  - Constrained
  - Partially constrained
  - Unconstrained

- **Interactivity**
  - Fixed
  - Fixed with illusion of choice
  - Interactive
<table>
<thead>
<tr>
<th>Interaction</th>
<th>Input</th>
<th>Human</th>
<th>Human + Automated tutor</th>
<th>Automated</th>
<th>Automated</th>
<th>Automated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Interactive</td>
<td>Interactive</td>
<td>Interactive</td>
<td>Fixed with illusion of choice</td>
<td>Fixed</td>
</tr>
<tr>
<td>Monologue</td>
<td>Constrained input</td>
<td>Registration form</td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td>Partially constrained input</td>
<td>Webinar. Usual web publication</td>
<td>Computer-assisted writing system</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td>Unconstrained input</td>
<td>Questions &amp; answers website [Yahoo Answers]</td>
<td>User-only initiative in chatbot [CSIEC...]</td>
<td>Questions-asking system where answer is not taken into account [ALICE-chan, Saybot]</td>
<td>Tutorial dialogue</td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td>Partially constrained input</td>
<td>Text adventure game and microworlds [FLAG, Spion, LINGO, MILT written]</td>
<td>Interactive unconstrained DIALOGUE SYSTEM [SPELL, Subarashii, DEAL, CSIEC, TLCTS, Façade...]</td>
<td>Free input, but ignored by system [Majestic]</td>
<td>Embedded dialogue (Cornillie et al 2013)</td>
<td></td>
</tr>
<tr>
<td>Dialogue</td>
<td>Partially constrained input</td>
<td>Utterance selection dialogue with multiple conversational paths [Let's Chat, Animated pedagogical agent]</td>
<td>Utterance selection with various correct options, but same reaction [Trace Effect]</td>
<td>Utterance selection with various correct options, but same reaction [Trace Effect]</td>
<td>Reading aloud prompted utterances in dialogue [Virtual Language Patient, FASOP]</td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td>Constrained input</td>
<td>Request more information form</td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*We excluded the possibility of having a fixed (or fixed with illusion of choice) interactional sequence with a human interlocutor (including automatically tutored), since we don’t know of any such interaction in computer-mediated environments. Yet, this possibility exists when acting, reading a script.

Focus-on-meaning: Successful communication is the main objective

Focus-on-form: Corrective feedback is the main objective
<table>
<thead>
<tr>
<th>Interaction</th>
<th>Input</th>
<th>Human</th>
<th>Human + Automated tutor</th>
<th>Automated</th>
<th>Automated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interlocutor</td>
<td></td>
<td></td>
<td>Interactivity</td>
<td>Interactivity</td>
</tr>
<tr>
<td>Monologue</td>
<td>Constrained input</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td>Partially constrained input</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td>Unconstrained input</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mono-initiative dialogue</td>
<td>Partially constrained input</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mono-initiative dialogue</td>
<td>Unconstrained input</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dialogue</td>
<td>Unconstrained input</td>
<td>2</td>
<td></td>
<td>82</td>
<td>1</td>
</tr>
<tr>
<td>Dialogue</td>
<td>Partially constrained input</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dialogue</td>
<td>Constrained input</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mono-initiative dialogue</td>
<td>Constrained input</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Focus-on-meaning**
Successful communication is the main objective

**Focus-on-form**
Corrective feedback is the main objective

- Text adventure games
  - [Spion, LINGO, MILT...]
- Mono-initiative chatbots
  - [Dave, Mike, CSIEC]
- Questions-asking systems
  - [ALICE-chan, Saybot]
- Computer-supported collaborative learning/Tutored SCMC
  - [MentorChat, Neckle...]
- Interactive & unconstrained DSLL
- Utterance selection
  - [Let’s Chat...]
- Reading aloud prompted utterances
  - [FASOP]
Why?

RATIONALE FOR DSLL
SCMC (chat) ⇒ L2 development

- SCMC ⇒ L2 oral proficiency

- SCMC ⇒ L2 oral proficiency, more than face-to-face
  (Sykes, 2005; Lin, 2015; although not confirmed by Ziegler, 2013)

- Why?
  - attention to form (noticing and feedback)
  - lower anxiety levels (Satar & Özdener, 2008)
DSLL → L2 development

- In foreign language teaching contexts: interactions in L2 often very rare

⇒ “Virtual immersion” (Ellis & Bogart, 2007; Fryer & Carpenter, 2006)

- Autonomous practice in a communicative, cultural, authentic and interactive task
  (Wachowicz & Scott, 1999; Fryer & Carpenter, 2006)
Available at any moment
Learner can go at his own pace (repeat, etc.)
Low anxiety environment
Potentially fully controllable learning environment
(e.g. feedback, learner modeling and adaptivity, motivational support)
What do we know?

Effectiveness Studies

https://databustersblog.files.wordpress.com/2014/10/control-groups.jpg
Types of studies on DSLL

- Experimental evaluation
- Case study
- System evaluation
- System
- Review

Studies published on DSLL

<table>
<thead>
<tr>
<th>Year</th>
<th>Experimental evaluation</th>
<th>Case study</th>
<th>System evaluation</th>
<th>System</th>
<th>Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Effectiveness studies

<table>
<thead>
<tr>
<th>Methodological Design</th>
<th>Perceived Usefulness</th>
<th>Impact on Motivation and Attitude</th>
<th>Impact on Language Knowledge and Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>14</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Pre-post + control</td>
<td></td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Pre-post + control with</td>
<td></td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>significant results</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experimental studies on DSLL’s impact on learning

- Vlugter, Knotta, McDonald & Hall (2009) 
  [Te Kaitito]: written DSLL tutorial in Māori
  o compares DSLL tutorial with in-class tutorial about Māori personal pronoun system

![Graph showing score progression over pre-test, 1st post-test, and 2nd post-test with two lines representing normal-tutorial system.]
Experimental studies on DSLL’s impact on learning

- Petersen (2010)
  [Sasha]: written DSLL in English providing corrective feedback (recast) and answering questions in a “spot the differences” task
  - compares Sasha vs. face-to-face interaction
    => impact on question construction
Experimental studies on DSLL’s impact on learning

- Wolska & Wilske (2010a)
  Written DSLL in German
  - compare **free production** vs. **constrained production** (gap filling) => impact on sentence construction and grammaticality judgement
  - N=15
    n=7
Experimental studies on DSLL’s impact on learning

- Wilske & Wolska (2011)
  - compares free production with either recasts or metalinguistic feedback, and constrained input => impact on sentence construction & grammaticality judgement
  - N=30 / n=9
How?

TECHNOLOGICAL PROCESS
LEARNING ENVIRONMENT

COMMUNICATION INTERFACE

Input mode(s)
- Written
- Spoken (ASR)
- Paralinguistic (vocal cues)
- Kinetic (face & body motion)
- Environmental

Output mode(s)
- Written
- Spoken (synthesized or prerecorded)
- Paralinguistic
- Kinetic (embodied agent)
- Environmental

INPUT PROCESSING

Surface operations
- Surface cleaning
- Surface reduction

Syntactic processing
- Part-of-Speech tagging
- Syntactic parsing

Natural language understanding
- Semantic parsing
- Dialogue act identification

OUTPUT GENERATION

Surface selection

Syntactic generation

Semantic generation

DIALOGUE MANAGEMENT

Reasoning and action selection
- Basic matching
- Graph (finite state)
- Frame (form-filling)
- Plan (AI planning, based on beliefs, desires and intentions detection)
DS: information state architecture

LEARNING ENVIRONMENT

COMMUNICATION INTERFACE

Input mode(s)
- Written
- Spoken (ASR)
- Paralinguistic (vocal cues)
- Kinetic (face & body motion)
- Environmental

Output mode(s)
- Written
- Spoken (synthesized or prerecorded)
- Paralinguistic
- Kinetic (embodied agent)
- Environmental

INPUT PROCESSING

Preprocessing
- Recognition errors corrections
- Spellcheck, etc.

Natural language understanding
- Keyword or concept spotting
- Grammar-based parsing
- Statistical parsing
- Shallow semantic parsing
→ Dialogue act identification

OUTPUT GENERATION

Surface selection
Syntactic generation
Semantic generation

DIALOGUE MANAGEMENT

Reasoning and action selection
- Basic matching
- Graph (finite state)
- Frame (form-filling)
- Plan (AI planning, based on beliefs, desires and intentions detection)
Initiative management in dialogue systems

**user-only initiative**
- question-answering systems
- text adventure games & microworld interactions

Main problems:
- References processing
- Information retrieval

**default user initiative**
- most chatbots

**mixed initiative**
- cooperative dialogue
- most task-oriented dialogue systems

Main problems:
- Natural language understanding
- Dialogue modeling and dialogue management
- Natural language generation

**default agent initiative**
- tutorial dialogue
  - most ITS

**system-only initiative**
- question-asking systems
- rigid tutorial dialogue

Main problems:
- Materials scripting
- Corrective feedback generation
Conclusions

- **What?**
  - Operational definition of DSLL based on systematic literature study:
    - “Dialogue activity with an automated agent, ideally with unconstrained input and interactive dialogue sequence”

- **Why?**
  - Insufficient experimental results to demonstrate effectiveness for language language
  - ⇒ Need for more effectiveness research

- **How?**
  - Advancements in dialogue systems haven’t yet been applied to DSLL:
    - natural language understanding
    - information state-based
    - data-driven / probabilistic models
  - ⇒ Need for more technological research
Thank you!

Dialogue systems for language learning:
state of the art and avenues for research
on task-based agents

Serge Bibauw
serge.bibauw@kuleuven.be

Thomas François
thomas.francois@uclouvain.be

Piet Desmet
piet.desmet@kuleuven.be