

Individualized feedback for lexical stress errors

Towards a CAPT system for French learners of German

Anjana Sofia Vakil

M.Sc. Thesis proposal
Language Science and Technology

Prepared under the supervision of
Prof. Dr. Bernd Möbius
Dr. Jürgen Trouvain



UNIVERSITÄT
DES
SAARLANDES

Saarland University
Department of Computational Linguistics & Phonetics

10 Never, 2014

Contents

1	Introduction	1
1.1	Context: The IFCASL project	1
1.2	Objectives	1
1.3	Thesis overview	2
2	Background and related work	3
2.1	Computer-Assisted Pronunciation Tutoring	3
2.1.1	Pronunciation in foreign language education	3
2.1.2	Computer-based and intelligent tutoring systems	3
2.1.3	Survey of existing CAPT systems	3
2.2	Towards CAPT for French learners of German	3
2.2.1	Phonetic and phonological comparison	3
2.2.2	Targeting lexical stress errors	3
2.3	Summary	4
3	System overview	5
3.1	Goal and architecture	5
3.2	Tools and technologies	5
3.3	User interface	5
4	Diagnosis of lexical stress errors	6
4.1	Related work	6
4.2	Automatic segmentation of nonnative speech	6
4.2.1	Segmentation via forced alignment	6
4.2.2	Evaluation of system accuracy	6
4.2.3	Coping with segmentation errors	6
4.3	Prosodic analysis	6
4.3.1	Pitch	6
4.3.2	Duration	7
4.3.3	Intensity	7
4.4	Comparison of native and nonnative speech	7
4.4.1	Using a single reference speaker	7
4.4.2	Using multiple reference speakers	7
4.4.3	Using no reference speaker?	7

4.5	Summary	7
5	Feedback on lexical stress errors	8
5.1	Related work	8
5.2	Visual feedback	9
5.2.1	Stylized text	9
5.2.2	Graphical representations of prosody	9
5.2.3	Visualizations of the speech signal	9
5.3	Auditory feedback	9
5.3.1	Enhanced reference utterance	9
5.3.2	Resynthesized learner speech	9
5.4	Alternative feedback types	9
5.4.1	Metalinguistic feedback	9
5.4.2	Interactive feedback	9
5.4.3	Implicit feedback	9
5.5	Summary	9
6	Conclusion and outlook	10
6.1	Thesis summary	10
6.2	Future work	10
	Bibliography	11

Introduction

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.1 Context: The IFCASL project

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.2 Objectives

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.3 Thesis overview

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Background and related work

2.1 Computer-Assisted Pronunciation Tutoring

2.1.1 Pronunciation in foreign language education

2.1.2 Computer-based and intelligent tutoring systems

2.1.3 Survey of existing CAPT systems

2.2 Towards CAPT for French learners of German

2.2.1 Phonetic and phonological comparison

Segments

Prosody

Other factors

2.2.2 Targeting lexical stress errors

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



Figure 2.1: Criteria for selecting errors to target in a CAPT system.

Frequency of production

Impact on intelligibility

Feasibility of automatic detection

2.3 Summary

System overview

3.1 Goal and architecture

3.2 Tools and technologies

3.3 User interface

Diagnosis of lexical stress errors

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

4.1 Related work

Lots of Nancy references

Duong et al., 2011

Probst et al., 2002

4.2 Automatic segmentation of nonnative speech

4.2.1 Segmentation via forced alignment

4.2.2 Evaluation of system accuracy

4.2.3 Coping with segmentation errors

Mesbahi et al., 2011, Orosanu et al., 2012

4.3 Prosodic analysis

4.3.1 Pitch

4.3.2 Duration

4.3.3 Intensity

4.4 Comparison of native and nonnative speech

Probst et al., 2002

4.4.1 Using a single reference speaker

Manually selecting a reference

Automatically selecting a reference

4.4.2 Using multiple reference speakers

4.4.3 Using no reference speaker?

Duong et al., 2011

4.5 Summary

Feedback on lexical stress errors

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

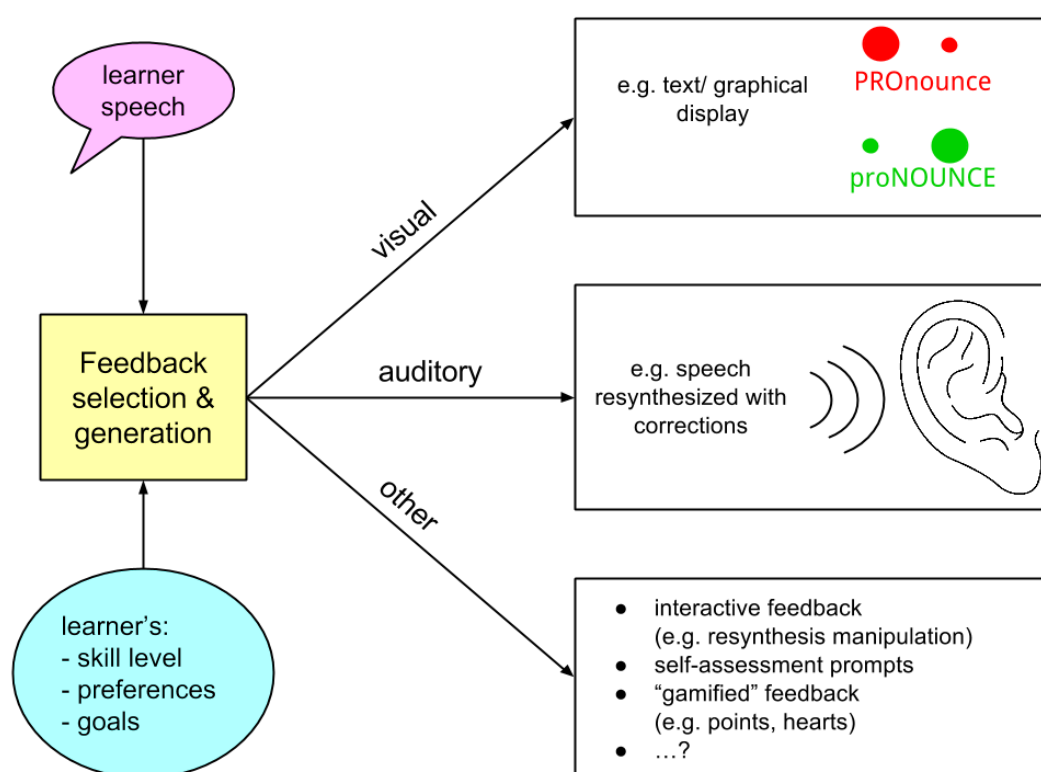


Figure 5.1: Delivery of prosody feedback in different modalities.

5.1 Related work

Sitaram et al., 2011 Bonneau and Colotte, 2011

5.2 Visual feedback

5.2.1 Stylized text

5.2.2 Graphical representations of prosody

5.2.3 Visualizations of the speech signal

5.3 Auditory feedback

5.3.1 Enhanced reference utterance

5.3.2 Resynthesized learner speech

5.4 Alternative feedback types

5.4.1 Metalinguistic feedback

5.4.2 Interactive feedback

5.4.3 Implicit feedback

5.5 Summary

Conclusion and outlook

6.1 Thesis summary

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6.2 Future work

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Bibliography

- Bonneau, Anne and Vincent Colotte (2011). “Automatic Feedback for L2 Prosody Learning”. In: *Speech and Language Technologies*. 1977. InTech.
- Duong, Minh, Jack Mostow, and Sunayana Sitaram (2011). “Two methods for assessing oral reading prosody”. In: *ACM Transactions on Speech and Language Processing* 7.212, pp. 1–22.
- Mesbahi, L, D Jouvét, A Bonneau, and D Fohr (2011). “Reliability of non-native speech automatic segmentation for prosodic feedback.” In: *SLaTE*.
- Orosanu, Luiza, Denis Jouvét, Dominique Fohr, Irina Illina, and Anne Bonneau (2012). “Combining criteria for the detection of incorrect entries of non-native speech in the context of foreign language learning”. In: *SLT 2012 - 4th IEEE Workshop on Spoken Language Technology*.
- Probst, Katharina, Yan Ke, and Maxine Eskenazi (2002). “Enhancing foreign language tutors – In search of the golden speaker”. In: *Speech Communication* 37.3-4, pp. 161–173.
- Sitaram, S, J Mostow, Y Li, A Weinstein, D Yen, and J Valeri (2011). “What visual feedback should a reading tutor give children on their oral reading prosody?” In: *SLaTE*.