

HQSYN16 - Task #3844

Task # 3669 (New): RA1a - Analysis and cataloguing of artifacts

Task # 3688 (Closed): Separation of some phonemes into distinct phones

Analyze utterances with phonetically incorrect phoneme

01.04.2016 10:26 - Tihelka Dan

Status:	Closed	Start date:	06.06.2016
Priority:	Normal	Due date:	20.11.2016
Assignee:	Skarnitzl Radek	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	RA1: Analysis of artifacts in synthetic speech		
Description			
We have added data synthesized in #3763 to wiki . Please look at them if it is what you have requested.			
Continue with analysis with the aim to answer the questions:			
<ul style="list-style-type: none">• must the phone variants be treated separately? I.e. must <i>[P]</i> speech units be used only where phonetic transcriptions says it belongs to, or can they be used in <i>[r]</i> positions under certain circumstances?• what are those circumstances? For example, is exchanging possible in some contexts?			
Related issues:			
Follows HQSYN16 - Task #3763: Synthesize utterances with phonetically incorre...		Closed	25.02.2016 03.06.2016

History

#1 - 01.04.2016 10:27 - Tihelka Dan

- Follows Task #3763: Synthesize utterances with phonetically incorrect phoneme realizations added

#2 - 21.04.2016 09:04 - Skarnitzl Radek

- Due date changed from 09.03.2016 to 02.05.2016

- Status changed from New to Assigned

#3 - 20.05.2016 11:42 - Skarnitzl Radek

- File *allophonic variants.docx* added

- Due date changed from 02.05.2016 to 20.05.2016

- Status changed from Assigned to Resolved

- % Done changed from 0 to 100

Allophonic variants

We performed a detailed auditory analysis of the synthesized phrases. Details are attached as a separate file, **allophonic variants.docx**, in Czech.

The results can be briefly summarized as follows:

1) [r] vs. [r=]

A normal [r] is very often realized as a flap (i.e., not a trill) in Czech. When such an [r] is inserted into words which should have [r=], quality is impaired - the target sound is very short and the absence of trilling is disturbing.

Sometimes there is an impression of an [r+vowel] sequence, rather than [r=].

2) [l] vs. [l=]

When a syllabic [l=] is used in contexts where a [l] would be followed by a vowel (i.e., *potlach* [potlax]), a glottal stop is usually inserted [potl=?ax], which is extremely disturbing.

Sometimes there is an impression of an [l+vowel] sequence, rather than [l=].

3) [n] vs. [N]

When alveolar [n] is used in pre-velar contexts, it is usually accompanied with an epenthetic schwa [len@ka].

When velar [N] is used in non-velar contexts, it is often audible and disturbing.

4) [x] vs. [G]

Here the phonetic voicing of the segment often did not correspond to the definition; none of the sequences was disturbing.

5) voiced ř [P] vs. voiceless ř [Q]

(Partially) voiceless realizations in contexts which should have a voiced realization is usually not disturbing, as ř is frequently (partially) devoiced in ordinary speech.

However, when the voicing of ř also changes the voicing of a neighbouring obstruent, the result is disturbing and may even lead to change of meaning (*dřímá* --> *třímá*).

When a voiced [P] is used in contexts which should be voiceless, the result is disturbing.

#4 - 24.05.2016 15:56 - Matoušek Jindřich

The auditory analysis will be performed for the utterances that were not synthesized (this was especially the case of [x] vs. [G]). So keeping this task unclosed, waiting until the utterances are re-synthesized ([#3763](#)).

A similar auditory analysis will be repeated for the voice Jan - see [#3922](#).

#5 - 01.06.2016 08:43 - Tihelka Dan

- Status changed from Resolved to Assigned

I have uploaded to [wiki](#) new files with synthesized utterances - now all should be correctly synthesized. Except the fixes of missing items and more variants of [ch] replacements, the remaining files contain the same data as were in the previous version (thus they do not have to be analysed again).

#6 - 21.06.2016 13:33 - Skarnitzl Radek

- File *allophonic variants2.docx* added

- Status changed from Assigned to Resolved

The second round of auditory analyses did not reveal any new information. The conclusions and recommendations from the previous phase hold:

- 1) the two versions of /l/ - plain [l] and syllabic [l̥] - should be kept separately, as the l -> l̥ interchange in prevocalic contexts almost always leads to synthesizing a glottal stop (e.g., [vLlak]).
- 2) the alveolar [n] and velar [ŋ] should be kept separately, the ŋ->n interchange usually yields an epenthetic schwa (e.g., [len@ka]).
- 3) the two versions of ř - voiced P and voiceless Q - should be kept separately
- 4) the allophones of /x/ - [x G h] - can be pooled together, without any marked intrusive effects.

#7 - 23.06.2016 16:48 - Tihelka Dan

- Status changed from Resolved to Assigned

The same as before, but now for *male* voice (Jan, spkr_AJ) - data put as attachment on [wiki](#).

#8 - 07.07.2016 13:30 - Skarnitzl Radek

- File *allophonic variantsJan.docx* added

- Status changed from Assigned to Resolved

Auditory analysis of the male speaker yielded more favourable results than for the female. Details are attached as a separate file, *allophonic variantsJan.docx*, in Czech, the main findings are summarized here:

- 1) based on the listening sample, the two versions of /r/ - plain [r] and syllabic [r̥] - could be pooled
- 2) the two versions of /l/ - plain [l] and syllabic [l̥] - could be pooled because the l -> l̥ interchange in prevocalic contexts never led to synthesizing a glottal stop (it almost always did with the female speaker).
- 3) the alveolar [n] and velar [ŋ] could also be pooled, the ŋ->n interchange led to an epenthetic schwa (e.g., [len@ka]) only once, and the effect was not very disturbing.
- 4) the two versions of ř - voiced P and voiceless Q - should be kept separately
- 5) the allophones of /x/ - [x G h] - can be pooled together, without any marked intrusive effects

In conclusion, it seems advantageous to specify the pooling or keeping apart of allophonic variants for every speaker in the database separately.

#9 - 08.08.2016 10:45 - Matoušek Jindřich

- Status changed from Resolved to Closed

#10 - 21.10.2016 13:42 - Matoušek Jindřich

- Due date changed from 17.08.2016 to 20.11.2016

- Status changed from Closed to Assigned

The same as before, but now for male voice (Stanislav, spkr_JS) and for female voice (Kateřina, spkr_SK) - data put as attachment on [wiki](#).

#11 - 05.12.2016 23:08 - Skarnitzl Radek

- File allophonic variantsKateřina.docx added

- File allophonic variantsStanislav.docx added

- Status changed from Assigned to Resolved

The results of the last set of analyses mostly follow the previous findings, and they are provided in the attached documents. Here the task is summarized and conclusions are drawn:

1) Regarding the **syllabic [r=] and [l=]**, it does not seem to be possible to generalize with respect to the speaker's sex. In the female voice, Katerina, the presence of syllabic [r=] led several times to the insertion of the glottal stop, e.g. *drozd* as [dr=lost]; this never happened with the male speaker, Stanislav. However, when [l=] was synthesized, a glottal stop was inserted in most of the cases in the male speaker's voice, Stanislav, and never with the female voice.

Conclusion: preferring to err on the side of caution, syllabic and non-syllabic sonorants - [l] and [l=], as well as [r] and [r=] - should be kept separately.

2) The interchange in synthesis of the **allophones of /n/ - [n] and [N]** - sometimes leads to the insertion of an epenthetic schwa (e.g., [len@ka]); the occurrence, as well as the salience of the schwa differs in individual speakers. Assuming that a nasal coming originally from the velar context would hardly be synthesized into an alveolar context in natural conditions (e.g., *vino* as [vi:No]), the answer to the question whether to keep the two contexts separately is not quite straightforward and depends on the richness of the segment inventory.

Conclusion: the epenthetic schwa is sometimes rather intrusive, but if the "cost" of keeping [n] and [N] as separate segments were to be too high, it is possibly to pool them.

3) The results are unequivocal with regard to the **Czech fricative trill /ř/** and its two allophones - the voiced [P] and voiceless [Q]. Interchanging one for the other typically leads to very intrusive effects.

Conclusion: the two allophones of /ř/ must be kept separately.

4) The effect of interchanging the allophones of the **velar fricative /x/** seems to be relatively minor, especially when it comes to [x] and its voiced counterpart [G]. The laryngeal [h] sometimes is intrusive, especially when it also triggers incorrect voicing assimilation; under normal conditions, this would most probably not happen, though.

Conclusion: the allophones of /x/ can be pooled together.

#12 - 11.12.2016 21:01 - Matoušek Jindřich

- Status changed from Resolved to Closed

Files

allophonic variants.docx	17.8 KB	20.05.2016	Skarnitzl Radek
allophonic variants2.docx	18.6 KB	21.06.2016	Skarnitzl Radek
allophonic variantsJan.docx	17.7 KB	07.07.2016	Skarnitzl Radek
allophonic variantsKateřina.docx	19 KB	05.12.2016	Skarnitzl Radek
allophonic variantsStanislav.docx	18.4 KB	05.12.2016	Skarnitzl Radek