The aim of this study is to investigate the perceptual salience of segmental differences between dialects, that is, of characteristics at the segmental level that distinguish one dialect from another (e.g. /ʌ/ vs. /œ/ in Grison German /meːl/ 'flour' vs. Zurich German /maːl/ 'flour'). In psycholinguistic and laboratory phonology studies, salience commonly refers to those elements in speech that are for listeners easily detectible (e.g., McGuire & Babel 2012; Silbert 2014, J. LabPhon). Sociolinguistic approaches to salience usually focus on the social meanings attributed to an element, which are assumed to further contribute to its conspicuousness (Auer 2014). Auditory (e.g. abrupt transitions to subsequent sounds), cognitive (e.g. the listener's phonological system and his experience with other varieties), and social factors (e.g. stigma attributed to a sound) are assumed to contribute to perceptual salience (Auer 2014, Linguistik Online). Acoustic/physiological salience (audibility) can be seen as the precondition that a feature acquires social-indexing meaning. Perceptual salience is considered to play a role in phonetic accommodation (e.g. Trudgill 1986, Dialects in Contact; Babel 2012, JPhon) and sound change (e.g. Ohala & Kawasaki 1984, Phonol. Yearbook) in the sense that more salient features are more likely to be adopted (or avoided) in linguistic interaction.

In the present study, the salience of segmental dialect differences is operationalized by investigating their contribution to speed and accuracy of dialect recognition in a forced-choice perception experiment. This approach is based, on the one hand, on the idea that more salient features contribute more to dialect recognition than less salient ones (MacLeod 2015, Ampersand); on the other hand, on evidence that phonetic-acoustic differences at the segmental level significantly predict dialect recognition in natural stimuli (e.g. Clopper & Pisoni 2004, JPhon). Thus, a shorter reaction time (as a measure of a faster decision process) will be interpreted as the result of a stronger link in the listener's mind between the perceived segment and the respective dialect.

The identification of Grison (GRG) and Zurich German (ZHG) stimuli by ZHG speakers is investigated. Among speakers of Swiss German, dialect is spoken in everyday situations and is considered high prestige. 48 mono- or bisyllabic words (the same for both dialects) were used and covered either none (12 words, e.g. 'lake' GRG /se/ vs. ZHG /se/), one (20 words, e.g. 'soup' GRG /suppe/ vs. ZHG /suppo/) or two (14 words, e.g. 'cook' GRG /kʰɔx/ vs. ZHG /xox/) differences at the segmental level as inferred from dialectological studies (Christen et al. 2010). Five vocalic (GRG/ZHG: /a/-/o/ /u/-/o/ /e/-/œ/ /o/-/ɔ/ /ɛ/-/o/) and 2 consonantal differences (/kʰ/-/k/) were included from which we hypothesised that GRG /kʰ/, /a/ and /u/ would be the most salient ones for ZHG listeners. These seven segmental dialect differences and the three identical vowels (/ɛ/, /o/, /u/) were distributed over the 48 words as equally as possible. Segmentally identical words were included as a baseline, as the effect of other cues (e.g. prosody, speech rate) cannot be excluded. Stimuli were spoken by 8 female native speakers of GRG, and 8 of ZHG (age range 19-24 years). Acoustic analysis (vowels, /n/-/n/) and auditory inspection (/kʰ/-/k/) confirmed the dialects significantly differed in the respective segments. The 768 stimuli were cut and amplitude-equalised. Participants heard the stimulus and had to decide whether the word was spoken in GRG or ZHG while reaction time (RT; from the onset of the stimulus) was measured. Stimuli were rotated over listeners so that each listener heard 192 stimuli (12 words per speaker; each word twice in GRG and ZHG) and pseudo-randomised to avoid the same speaker or the same word appeared in succession. 18 native monolingual speakers of ZHG participated in the experiment (7 males, 11 females, 19-32 years old).

Fig.1 shows that accuracy was generally higher in ZHG than GRG stimuli, and higher the more segmental dialect differences occurred in the stimulus. A generalised linear mixed effect model (GLMM) with accuracy as the dependent variable, dialect and number of segmental differences as fixed factors, and word, speaker and listener as random factors showed a significant effect of number of differences and dialect. An interaction between number of differences and dialect likely came about because of the listeners' tendency to ascribe a stimulus without a segmental dialect difference to their own dialect. Fig. 2 shows how accuracy varies with segmental difference in words with 0 or 1 segmental dialect differences. A second GLMM with accuracy as the dependent variable, the same random factors as above, and dialect and segmental difference (7 levels) as fixed factors showed a significant main effect of segmental difference and a significant interaction between dialect and segmental difference. Tukey tests confirmed that all GRG stimuli with a segmental dialect difference
were significantly better recognised than those without. Accuracy was highest in stimuli with /k/ and lowest in stimuli with /a/. Among ZHG stimuli only those containing /k(\text{x})/ showed a significantly higher accuracy than those without a segmental dialect difference.

RTs of correct answers over 2500 ms were removed from the data set (53/2682 tokens = 2.0%). From this trimmed data set, RTs more than 2.5sd from the mean were considered as outliers and removed (147/2735 = 5.4%). Overall, RTs for correct answers were longest in GRG stimuli without a segmental dialect difference (Fig. 3). A linear mixed effect model with RT as the dependent variable, number of segmental differences and dialect as fixed factors, and random factors speaker, word, and listener showed a significant interaction between the fixed factors. Tukey tests confirmed that RTs were longer in GRG stimuli without segmental dialect differences than in all other contexts. A second model with dialect and segmental difference (7 levels) as fixed and the above mentioned random factors showed an interaction between the fixed factors. Tukey tests showed that RTs to GRG stimuli were shortest in words containing /k/, /a/, /ɔ/, longest in segmentally identical words, and intermediate in words containing /æ/, /ɛ:/, /ʊ/ (see Fig. 4). Among the ZHG stimuli only the differences between /k/-/a/ and /k/-/ɔ/ were significant.

The results suggest that the presence of a segmental dialect difference affects accuracy and speed of dialect recognition. The effect of number and type of segmental dialect difference on RT and accuracy was more marked in GRG stimuli, though listeners relied more strongly on segmental dialect differences when identifying a dialect different from their own. Although other, especially suprasegmental features, are likely to have influenced listeners’ decisions, the number and type of segmental dialect features consistently affected speed and accuracy of responses. According to the results, the most salient GRG features for ZHG listeners were /k/ > /æ/, /a/, /ɛ/. An online questionnaire about explicit knowledge of typical GRG pronunciation features (24 independent speakers of ZHG; non-linguists) revealed a similar ranking: /k/ (20 mentions) > /æ/ (9 mentions), indicating a relationship between implicit and explicit knowledge. While /k(\text{b})/-/k(\text{x})/ as a discrete difference between the ZHG and GRG certainly meets the auditory salience criteria, word-final /\text{v/-a}/ can be considered to be auditorily less salient than other vowel pairs since it occurs only in unstressed syllables and shows considerable acoustic within-dialect variability. Importantly, however, /æ/ is locally restricted, while /a/, /ɛ/, /ɔ/ and /ʊ/ are wide-spread among Swiss German dialects. This aspect might have contributed to the participants’ explicit knowledge and could explain the different results for implicit and explicit knowledge. Further research is needed to understand how acoustic distance, the listener’s grammar and the listener’s linguistic experience contribute to perceptual salience.