Articulatory overlap as a function of voicing in French and German consonant clusters

Lasse Bombien[a)] and Philip Hoole
Institut für Phonetik und Sprachverarbeitung, Ludwig Maximilians-Universität München, Schellingstraße 3, D-80799 Munich, Germany

(Received 6 March 2012; revised 19 April 2013; accepted 26 April 2013)

The effects of laryngeal specification on the timing of supra-laryngeal articulations have so far received little attention. Previous research has shown that German—but not French—mixed-voicing clusters are produced with less articulatory overlap than phonologically fully voiced clusters. Articulatory and acoustic data of labial and velar stops as simple onsets and in stop + /l/ clusters are examined to probe the causes for this cross-linguistic difference in the light of the different voicing implementations of French and German. The absence of overlap in German mixed-voicing clusters is attributed to the requirement of a time slot for the stop's aspiration phase. Since French does not commonly have aspirated stops, French clusters are expected to pattern with the voiced German clusters. The results confirm that voicing patterns established for simple onsets in the literature in terms of voice onset time of both German and French also obtain in clusters. Furthermore, the data show that contrary to the expectations French clusters pattern with German mixed-voicing clusters. This low degree of overlap in both voiceless and voiced French clusters indicates that overlap is restricted by aerodynamic requirements which result from the implementations of the voicing contrast.
 [http://dx.doi.org/10.1121/1.4807510]

PACS number(s): 43.70.Kv, 43.70.Bk [AL] Pages: 539–550

I. INTRODUCTION

The coordination of sequences of supra-laryngeal articulations with respect to laryngeal specification is an area of speech production research which so far has received only limited attention and is far from being understood. This becomes especially clear when considering data as reported, e.g., by Hoole *et al.* (2009) in a cross-linguistic study of French and German consonant clusters using electromagnetic articulometry (EMA). This study found that the laryngeal contrast conditions articulatory overlap in word-initial consonant clusters in German between the first consonant (C_1) and the second (C_2). More specifically, sequences such as /gl/ and /bl/ where C_1 is underlyingly, i.e., according to its phonological specification, voiced (harmonically voiced clusters), overlap to a higher degree than /kl/ and /pl/, where C_1 is voiceless (mixed-voicing clusters). This finding is in agreement with reports on Georgian consonant clusters: Chitoran *et al.* (2002) point out that clusters with complex laryngeal specifications (i.e., mixed-voicing clusters) are produced with little overlap. In French clusters, on the other hand, harmonically voiced clusters and mixed voicing clusters did not seem to differ with regard to articulatory overlap. The present study aims at shedding further light on this finding by linking this cross-linguistic difference to another better known fact: French and German are considered to differ in the means of implementing voicing contrasts in initial stop consonants. French accomplishes the contrast by the use of (true) voicing whereas German employs aspiration: phonologically voiced stops are usually not voiced while phonologically voiceless stops are indeed voiceless but also post-aspirated. Traditionally, aspiration does not play a role in French. Voice-onset time (VOT), the duration from stop release to onset of phonation, has commonly been employed to characterize this difference; see details below. In this paper, synchronous articulatory (EMA) and acoustic data are evaluated to make inferences on the influence of laryngeal specifications on the coordination of oral articulations.

A. Voice onset time and the voicing contrast in French and German word-initial stops

Lisker and Abramson (1964) state that characterizing voicing contrasts in stops can successfully be accomplished using VOT. VOT is the temporal distance from the release of the stop in question to the onset of voicing. This distance can be (a) positive (long and short lags), for example, in voiceless aspirated stops where voicing starts after the release of the stop, (b) negative, for example, in voiced/prevoiced stops, where voice onset is prior to the release of the stop, and (c) zero in voiceless unaspirated stops. Other measures or acoustic properties on their own fail to account for the various different mechanisms that the world's languages employ to create the voicing contrast. In a condensed view of English, true physiological voicing, i.e., "the presence of a glottal buzz" (Lisker and Abramson, 1964, p. 384) or its absence, reliably separates word-medial and final /b d g/ from /p t k/ but it fails word-initially since there both groups are generally produced without vocal fold vibration. Aspiration, on the other hand, distinguishes /p t k/ from /b d g/ in word-initial and medial position but is less successful word-finally since there is often no aspiration in /p t k/ and not even an audible release in /b d g/. Lisker and Abramson therefore conclude for English that neither voicing nor aspiration alone can account for the phonological voicing contrast.

a)Author to whom correspondence should be addressed. Electronic mail: lasse@phonetik.uni-muenchen.de

According to Maddieson (2009) (among others) languages with bi-modal voicing contrasts typically pattern into having either a prevoiced vs short-lag opposition as for example French (also Caramazza and Yeni-Komshian, 1974) and Spanish or a short-lag vs long-lag opposition as for example English and German.

Few sources are available for inter-gestural coordination in (voiceless) stop sounds for French and German. In their fiberscopic analysis, Benguerel *et al.* (1978) state that in French the glottal devoicing gesture is timed with the occlusion such that the glottal abduction-adduction movement starts at the same time as the oral occlusion and ends at the release or shortly after that. In recent work on German, Hoole (2006) used fiberscopic transillumination to analyze laryngeal-oral coordination in a large set of consonants both as singletons and in clusters in word initial position. The set included both /p/ as well as /pl/. In his data, glottal abduction starts shortly after the onset of oral occlusion with peak glottal opening (PGO) in close vicinity to the release of the oral occlusion. This means that a considerable portion of the glottal aperture remains after the occlusion which is quite contrary to the patterning described for French data.

The study of Klatt (1975) is of immediate importance since unlike most other studies it also deals with VOT in English consonant clusters. Among others, it covers the clusters under analysis in this study: /kl gl/ and /pl bl/ as well as the corresponding singleton stops. The most general finding is, of course, that voiceless stops have considerably longer VOT than voiced stops thus confirming the typical Germanic dichotomy of short lag vs long lag. Furthermore, three observations are relevant to the study at hand. The first is the universal (Maddieson, 1997) fact that place of articulation has an effect on VOT with shorter values for labial than for lingual stops. This has been demonstrated in quite a number of studies for a range of languages before Klatt and afterwards, e.g., Lisker and Abramson (1964, 1967); Weismer (1979); Crystal and House (1988); Docherty (1992); Nearey and Rochet (1994); Cho and Ladefoged (1999); Hoole (2006). Some of these works, but not all, also find that velars have longer VOT than apicals. Second, VOT has cross-linguistically been found to vary as a function of the following vowel's height (e.g., Fischer-Jørgensen, 1972). After high vowels, VOT is generally longer than after low vowels. Both phenomena have been attributed to the fact that VOT lengthens as a function of constriction degree of the vocal tract after stop release. In other words, a slow depletion of supraglottal air pressure due to a slow opening of the vocal tract impedes the onset of voicing. Stevens (2000) and Cho and Ladefoged (1999) point out that specific aerodynamic factors may also play a role in terms of the Bernoulli forces between the articulators forming the constriction: Velars have a longer and narrower constriction than, e.g., bilabials and consequently larger Bernoulli forces prolong the critical constriction. Similarly, high vowels, e.g., /i/, must in themselves be regarded as constrictions unlike low vowels for which the tongue and jaw lowering cause the vocal tract to assume its most unconstricted state. Klatt's third observation is related to the second: in stop + /l/ clusters VOT tends to be longer than in sequences of stop + low vowel. Similarly to stop + high vowel sequences, greater VOT in stop + /l/ is attributed to a higher degree of constriction than in stop + low vowel sequences.

B. Research aims

The following paragraphs summarize the research aims addressed in this study. The central concern of this study is to investigate temporal coordination of the articulation of initial consonant clusters in harmonic vs mixed voicing clusters in French and German. In order to make a connection to how these languages implement the voicing contrast in initial stops, the first part basically aims at establishing that the data adhere to the voicing patterns for French and German as they have previously been reported. Of greater interest, however, is whether these patterns also pertain in clusters. Based on the literature the following results are expected for VOT and the occlusion duration.

E1: German stop voicing contrast is realized by a short-lag/long-lag opposition in VOT whereas in French the opposition is one of prevoiced/short-lag. This also applies to clusters. Occlusion duration should be longer in voiceless (fortis) stops. /k/ has shorter occlusion than /p/ both in singletons and in clusters. /p/ has shorter VOT than /k/ both in singletons and in clusters.

Since articulatory data are available, this study also analyzes articulatory constriction plateau durations (see Sec. II C) in order to investigate possible effects of voicing, place of articulation and language on properties of intra-gestural coordination.

E2: It is expected that C_1 constriction and occlusion duration show basically similar patterns as a function of voicing, place of articulation, and language.

The present study will also test whether the difference in sensitivity to voicing observed by Hoole *et al.* (2009) for cluster timing in French vs German holds across a larger number of speakers. Previous findings give rise to the following expectations for inter-gestural coordination.

E3: Based on the previous findings, larger overlap is expected in German harmonically voiced clusters than in mixed voicing clusters. Overlap in French clusters should be insensitive to voicing.

While there are some indications that overlap might vary as a function of stop place of articulation, it is doubtlessly problematic to compare overlap in absolute terms when different articulators are involved (here lips vs tongue back). We will thus simply report place of articulation results because they form an integral part of the statistical design, but refrain from further interpretation.

A central concern of this study is how the difference in sensitivity to voicing observed for clusters in German and French is temporally implemented. Two scenarios for cluster timing in French emerge (see Fig. 1).

(1) Overlap in French stop + /l/ clusters is more like overlap in *harmonically voiced* clusters in German.
(2) Overlap in French stop + /l/ clusters is more like overlap in *mixed voicing* clusters in German.

Considering the case of German, a reasonable assumption would be that articulatory overlap is less in the context

FIG. 1. Schematic overlap pattern of constriction plateaus for German stop + sonorant clusters. Harmonically voiced clusters have overlapping plateaus, mixed voicing clusters have lags between the plateaus. Right hand annotations indicate the scenarios for timing in French.

of a voiceless stop in order to temporally accommodate the glottal gesture or more precisely the aspiration phase which is due to glottal timing. A perceptual motivation behind this might be the more continuous sonority[1] modulation (Ohala, 1992; Ohala and Kawasaki-Fukumori, 1997): As it is, i.e., large lag between the stop and /l/, a sonority profile of the following order is likely to emerge: voiceless stop–aspiration–voiceless lateral (fricative?)–voiced lateral–vowel. This sequence could be expressed as a series of uniform rises in sonority and also captures the fact that sonorant devoicing is often found to be partial (e.g., Tsuchida and Cohn, 2000, for English). Increasing overlap might lead to a fully devoiced lateral and, importantly, a voiceless transition from the lateral into the vowel both of which can be considered perceptually adverse: Presumably, a voiceless transition from the lateral into the vowel provides less clear information on the articulatory trajectories. Furthermore, the recoverability of the transition from the stop to the lateral would be endangered since crucial burst characteristics are masked by the lateral. Given this consideration, one could assume that the greater lag in the voiceless case is the result of a perceptually motivated rightward shift of the lateral.

E4: Since in French the glottal gesture must be timed differently in order to account for the fact that less or no aspiration occurs, there is no need to shift the lateral to the right since it is at no risk of undergoing total devoicing. It would therefore seem appropriate to assume that French clusters are timed more like voiced German clusters, i.e., with more overlap than voiceless German clusters (second scenario).

II. METHOD

A. Speakers and speech material

The material analyzed here is a subset of the same data corpus investigated by Hoole *et al.* (2009). However, the number of speakers has been increased by two for each language. Accordingly, five speakers each of French and German were recorded by means of EMA. The test corpora [French (FR) and German (DE)] are part of a larger project and were designed to contain all possible word onsets of the respective language. For each word onset, two words were selected, one with a low back vowel the other with a high front vowel following the onset (e.g., "Bad" [ba:t] (bath) and "Biest" [bi:st] (biest). This study uses only subsets of these corpora containing the simple onsets /b/, /g/, /p/, and /k/ as well as the same consonants forming complex onsets with a following /l/. These subsets are presented in Table I. The choice of material was based on previous findings that the /kl/ cluster exhibits the most stable coordination patterns of the clusters analyzed and because it has a fully-voiced counterpart /gl/. /pl/ and /bl/ were chosen because they present the only other pair of clusters with this voicing contrast in German that does not involve velic activity. The target words were embedded in carrier sentences which had three slots for the target words:

DE: Ich sage wieder «*word*#1» oder «*word*#2» oder «*word*#3».
I say again «word#1» *or* «word#2» *or* «word#3».
FR: Je vois «*word*#1» ou «*word*#2» ou «*word*#3».
I see «word#1» *or* «word#2» *or* «word#3».

The randomization routine ensured that this study's target words were distributed equally between the first and the second position (not the third).

For the two more recently recorded speakers of French, the carrier phrase and the material were altered to enable additional analyses which require more control of the coda of the target word. Variation of vowel height was dropped in this corpus such that the target words only contained low vowels. These modifications do not affect the current study since its subject matter does not involve the coda. The altered carrier phrase had only one target slot.

FR: Je vois «*word*» à l'écran.
I see «word» *on the screen.*

B. Data acquisition

Acoustic and electromagnetographic data were obtained simultaneously. The AG500 EMA system (Carstens Medizinelektronik) was used with sensor coils attached to the upper and lower lip, to the jaw on the gum just below the lower incisors and to the tongue (tip, mid, and back). The

TABLE I. Material for the analysis of voicing in clusters in French and German stop + /l/ clusters.

	French			
	Voiced		Voiceless	
Onset	high vowel	low vowel	high vowel	low vowel
velar				
simplex	—	gâte	kif	cap
complex	glisse	glace	clique	claque
labial				
simplex	bique	bac	pic	pâte
complex	blini	blatte	plisse	plaque
	German			
velar				
simplex	gib	gab	kies	kahl
complex	glied	glas	klean	klag
labial				
simplex	biest	bad	piep	pack
complex	blieb	blatt	plitsch	plan

tongue back sensor was placed as far back on the tongue as the subjects tolerated, the tip sensor about 1 cm behind the tip because locations closer to the tip are often felt by subjects to interfere unduly with articulation. Additionally, sensor coils were placed behind the ears on sections of skin that were inert to speech movements, on the bridge of the nose and on the gum above the upper incisors. The latter four sensors were used for head movement correction since their positions can be assumed to remain stable. Positions were calculated from the raw EMA amplitudes using the TAPAD (Three-dimensional Artikulographic Position and Align Determination) algorithm under MATLAB (see Hoole and Zierdt, 2010, for further details of the EMA processing). All EMA data were sampled at 200 Hz, rotated to the occlusal plane with the origin approximately at the incisors. The movement data was filtered using a Kaiser filter design with a pass-band edge of 20 Hz and a stop-band edge of 30 Hz for all sensors except the tongue tip, for which the edge frequencies were 40 and 50 Hz, respectively. Velocities were calculated by convolving the same filters with a differentiation kernel. Audio data were obtained using a Sennheiser microphone (MKH 50 P48) and recorded by means of a Sony multichannel DAT recorder (PC208Ax) at a rate of 24 000 Hz. Noise induced by the magnetic field emitters was filtered out. The speech material was presented to the speakers 10 times in randomized order on a computer screen.

C. Extraction of temporal parameters

Figure 2 schematically shows the positioning of articulatory landmarks by reference to an arbitrary articulator's position and its absolute velocity over time. Maximum onset velocity (2), maximum constriction (4), and maximum offset velocity (6) are easily detectable from the respective signal. The other landmarks, onset and offset of the gesture (1, 7) and begin and end of the constriction plateau (3, 5), are interpolated values and represent the 20% threshold of the difference between two adjacent extrema in the velocity signal, e.g., begin of constriction plateau (3) is positioned at the 20% threshold between maximum onset velocity and maximum constriction as detected in the respective signal. The 20% threshold method was preferred over actual zero crossings (or local minima for tangential velocity) because sometimes more than one zero crossing can occur during and after the target phase. The use of a velocity threshold also serves to distinguish directed actions of an articulator from coarticulatory movements. Given the phonetic material in the present study, three different articulatory trajectories were subjected to this procedure: The time function of the vertical position of the tongue back sensor and its velocity signal for the dorsal consonants /k, g/; the Euclidean distance (d) between the sensors attached to the upper lip (UL) and lower lip (LL) and the derived velocity as a measure of lip aperture for /p, b/ [two-dimensional, anterior-posterior (y) and vertical (z) dimension; $d(UL, LL) = \sqrt{(UL_y - LL_y)^2 + (UL_z - LL_z)^2}$]; the vertical position (z) of the tongue tip sensor using the tangential velocity (v_t) for /l/ ($vt = \sqrt{v_y^2 + v_z^2}$), anterior-posterior ($y$) and vertical ($z$) dimension. Figure 3 illustrates this procedure for a consonant sequence. Especially in the case of tongue tip movements sometimes several points matched the velocity criterion for plateau landmarks due to variability in the tangential velocity. In such cases we selected the point which resulted in a pattern most consistent with the cases where the procedure was unambiguous.

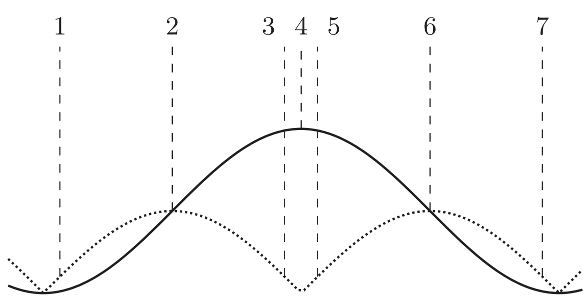

FIG. 2. Schematic display of landmark positioning. (1) onset of gesture; (2) maximum velocity in onset; (3) beginning of constriction plateau; (4) maximum constriction; (5) end of constriction plateau; (6) maximum velocity in offset; (7) offset of gesture. (1), (3), (5), (7) positioned using 20% threshold, see text. Solid line: (arbitrary) articulatory trajectory; dotted line: (absolute) velocity.

The usage of different measurement techniques (positions of tongue sensors vs lip aperture) and of different types of velocities (vertical for tongue back vs tangential for tongue tip) result in some caveats when it comes to the comparison of different articulators. For example, constriction plateaus do not have the same meaning when derived from movements of an articulator against a hard structure (as in apical and dorsal stops) compared to plateau durations derived from the movement of two soft articulators against each other (as in lip aperture). In the case of lip aperture movement may continue even though a complete closure has already been achieved. Furthermore, it is well known (e.g., Mooshammer *et al.*, 1995) that there is typically also a horizontal component in the tongue dorsum trajectory of velar stops and which can only be captured along with the vertical component when using tangential velocities and not unidimensional velocities. However, we believe that for a number of reasons this does not threaten the validity of the current study: First of all, in our data the main opening/closing movement of the tongue dorsum was indeed in the vertical dimension and especially in complex onsets the presence of the alveolar lateral reduced horizontal movement. Furthermore, while C_1 can be either a velar or a bilabial stop and different methods were employed for landmark detection, C_2 (/l/) landmarks were always detected with the same method. Most importantly, however, this study's focus does not lie on place of articulation but on coordination differences due to language and the voicing contrast.

For consonant clusters, a measure of articulatory overlap was calculated: Plateau overlap is the extent to which the two consonants' plateaus overlap, i.e., the period between constriction plateau onset of the second consonant (C_2) and the constriction plateau offset of the first consonant (C_1), i.e., $B5$–$T3$ in Fig. 3. Positive values indicate that the plateaus indeed overlap, negative values indicate a lag between the plateaus.

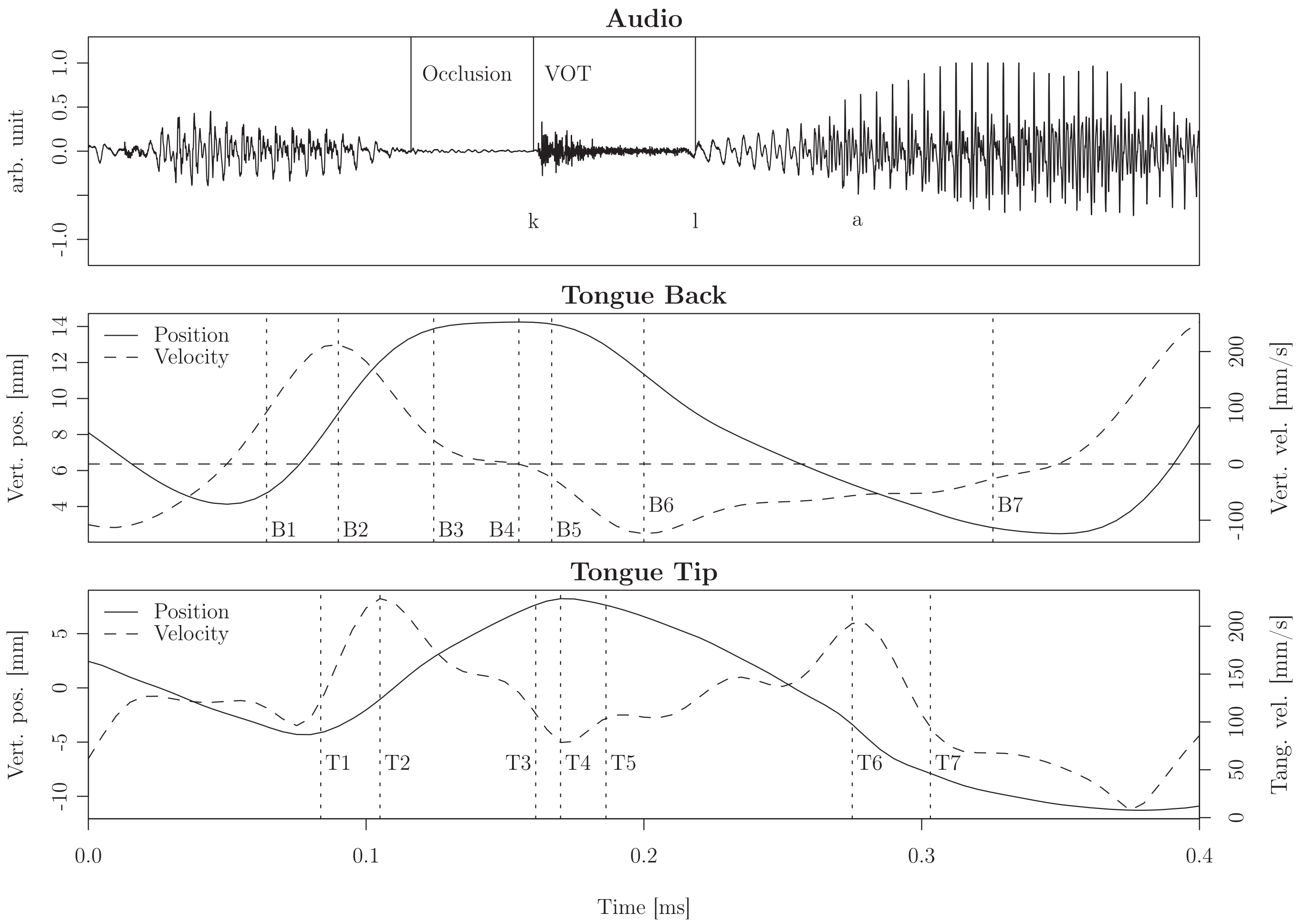

FIG. 3. Extraction of temporal parameters in a case of the word "klag" (/kla:k/). Upper panel illustrates the segmentation of occlusion duration and VOT and contains a roughly aligned transcription. The lower panels show the vertical position of tongue back (middle) and tongue tip (bottom) sensors in solid lines (axis labels on the left) and the velocity in dashed lines (axis labels on the right). For the tongue tip, the tangential velocity is used for landmark detection, therefore no horizontal zero line is present as opposed to the signed vertical tongue back velocity. Vertical lines in the lower panels correspond to articulatory landmarks labeled according to articulator (T = tongue tip, B = tongue back) and landmark number as defined in Fig. 2.

Normalization of overlap was not carried out because of the lack of a reliable normalization parameter. Typical parameters such as duration of constriction formation or plateau of C_1 are likely to be subject to strong variation due to effects of place of articulation (e.g., Byrd, 1993; Maddieson, 1997) and of voicing (see Fuchs, 2005). C_2 plateaus also proved to be unreliable, see Sec. III C 1.

The acoustical measures, C_1 occlusion duration and voice onset time, were defined as follows: Occlusion duration starts at the beginning of the occlusion as determined from waveform and spectrogram. It ends at the stop's release. VOT is here defined to start at occlusion offset. It ends at the onset of periodicity following the stop burst, cf. Fig. 3. This interval may be zero but not negative. Phonologically voiced stops in German vs French differ in that French stops are fully voiced whereas German stops usually are not. A measure of voice lead/voicing during closure/negative VOT would be appropriate to capture this difference. However, no glottal abduction-adduction gesture is involved in French or German voiced stops and consequently no inferences can be made about the coordination of laryngeal and oral articulations. A measure of voice lead was therefore not included in this study.

D. Statistics

The analysis of effects on VOT due to vowel quality as described above is not pursued here because the variation of vowel height was not recorded for two of the French speakers. Instead the data are pooled in this regard. Should any statistical blurring result from this, it should only enhance the power of significances found in other regards. Since VOT differences between single stops and stop + /l/ clusters are only reported to emerge in low vowel contexts the matter of increased VOT in stop + /l/ clusters is dropped here as well. Linear mixed effect models were calculated using speakers as random factors. The rate of speech can possibly change during the course of the experiment. In order to account for such effects, the repetition number was included as a random factor as well in a version of the model. All cases for which the variance explained by repetition number exceeds 0 will be pointed out. Table II lists all predictors (independent variables) and variables used in the statistical models. For consistency, the term C_1 plateau duration is also used for simple onsets. It always denotes the plateau duration of the initial stop.

A potential problem with mixed models is the determination of the degrees of freedom in the denominator and

TABLE II. Predictors and dependent variables used in the statistics in Sec. II.

Predictor	Description
LANG	Language: French or German (FR/DE)
PLAC	Place of articulation: (bi-)labial or velar (L/V)
VOX	Voicing: voiced or voiceless (phonologically) (+V/−V)
COMP	Complexity: complex or simple onsets (C/S)
SPK	Speakers
REP	Repetition/block number
Variable	Description
VOT	Voice onset time in ms
OCC	Occlusion of the stop in ms (as measured in the acoustics)
C1p	C_1 plateau duration in ms
C2p	C_2 plateau duration in ms
POVER	Plateau overlap in ms

hence the calculation of p values. Following Reubold *et al.* (2010), the degrees of freedom in the denominator (df_2) were arbitrarily set to 60 to avoid significances for only small changes in the F value. An alternative and less conservative approach by Baayen (2008) (p. 241 ff) estimates as follows: $df_2 = n - k - 1$, where n is the number of observations and k the number of degrees of freedom. This approach yields values between 363 and 1469 in the present data. High values in the denominator can lead to significances for only small changes of the F value, e.g., the F values for $\alpha = 0.01$ are $F[1, 1469] = 6.65$ but $F[1,60] = 7.07$, i.e., a change of 0.42 in the F value. By setting df_2 to 60, the statistical results can be regarded as more conservative. The selected degrees of freedom result in the following F value thresholds: $F = 4.00$ for $\alpha = 0.05$; $F = 7.07$ for $\alpha = 0.01$; $F = 11.97$ for $\alpha = 0.001$.

III. RESULTS

The results of this study are listed in the order of the expectations outlined in Sec. I B. In a first step we describe the patterns of VOT and occlusion to set the scene for the subsequent presentation of intra-gestural and inter-gestural coordination.

A. VOT patterns and C_1 occlusion duration in single and complex onsets

The aim of this section is to address the issues raised in E1 and to investigate the effects of voicing, language, place of articulation, and onset complexity on the occlusion duration and VOT as measured in the acoustics. An overview of the results is given in Table III and in Fig. 4.

1. Voice onset time

The main issue in addressing voice onset time here is to test whether complex onsets have longer VOT than simple onsets. The data for French voiced stops are, as mentioned above, not meaningful in this context: With very few exceptions French voiced stops had voicing during closure and non-zero VOT values are here due to short voicing interruptions at the release. In order to exclude these data, a first model was fitted only for voiceless/mixed-voicing onsets as a function of language and place of articulation.

The expected language specific implementations of the voicing contrast are reflected in that VOT is 43 ± 7 ms shorter in French stops than in German stops ($F[1,60] = 39.1$, $p < 0.001$). Velars have generally 23 ± 1 ms longer VOT than labials ($F[1,60] = 347.7$, $p < 0.001$) which is also in line with previous findings as reported in Sec. I. A barely significant effect of complexity ($F[1,60] = 4.4$, $p < 0.05$) accounts for 2.5 ± 1 ms shorter VOT in complex onsets than in simple onsets which is contrary to the expectations but complexity is also involved in two interactions: Language and complexity ($F[1,60] = 60.4$, $p < 0.001$) interact in such a way that in German VOT is 11 ± 2 ms shorter in complex onsets than in simple onsets ($F[1,60] = 27.4$, $p < 0.001$) while in French VOT is 6 ± 2 ms longer in complex than in simple onsets ($F[1,60] = 10.1$, $p < 0.01$). The interaction of place and complexity ($F[1,60] = 60.4$, $p < 0.001$) is due to a tendency for velars to have shorter VOT and labials to have longer VOT in complex than in simple onsets.

According to a model fitted to VOT data of voiced German onsets as a function of complexity and place VOT is 12 ± 1 ms longer in velars than in labials ($F[1,60] = 169.8$, $p < 0.001$). Complexity adds very little to the duration of VOT (2 ± 1 ms; $F[1,60] = 4.1$, $p < 0.05$). The main impact of complexity is however only found in labials (place × complexity: $F[1,60] = 5.2$, $p < 0.05$) where complex onsets have 4 ± 1 ms longer VOT than simple onsets ($F[1,60] = 18.5$, $p < 0.001$). Effects of complexity are rather weak in comparison to effects due to place of articulation.

2. Occlusion duration

Turning to occlusion, a mixed model was fitted to occlusion as a function of place, voicing, language, and complexity. Place of articulation accounts for an average of 16 ± 1 ms shorter occlusions in velars than in labials ($F[1,60] = 437.9$,

TABLE III. Summary of effects on both complex and simplex onsets. Interactions are only presented when they contribute crucially to the understanding of the data. Significance codes: *** ($p < 0.001$); ** ($p < 0.01$); * ($p < 0.05$); n.s. (not significant).

	Predictor			
Measure	Place (Vel/Lab)	Voicing (+V/−V)	Language (DE/FR)	Complexity (C/S)
VOT (−V only)	Vel > Lab ***	—	DE > FR ***	FR: C < S ** DE: C > S ***
VOT (+V DE)	Vel > Lab ***	—	—	Lab: C > S ***
Occlusion	Lab > Vel *** DE > FR	−V > +V *** FR, Lab, C	FR > DE ** Vel	S > C ***
C_1 plateau	Lab < Vel ***	−V > +V *** FR, Lab	n.s.	S > C ***

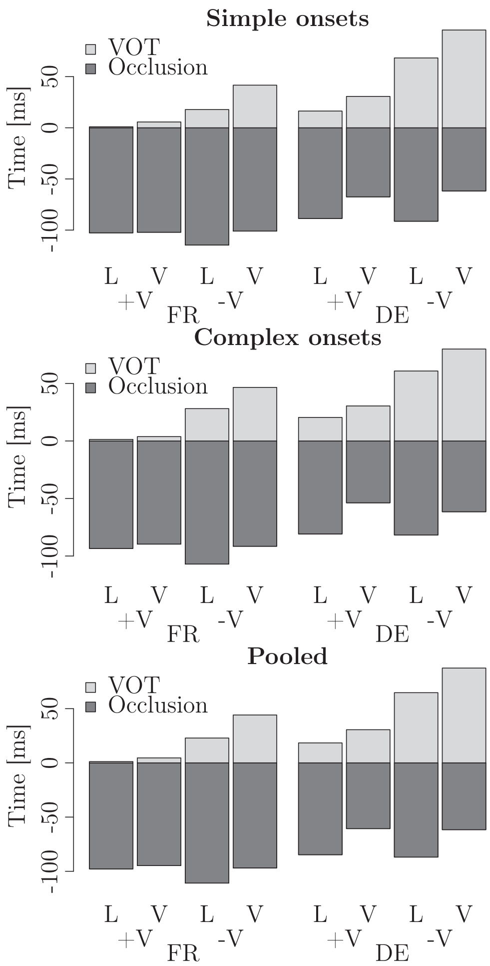

FIG. 4. Mean durations of acoustical occlusion and VOT in simple (upper panel) and complex (middle panel) onsets as a function of language [French (FR) vs German (DE)], voicing (+V vs −V) and place [labial (L) vs velar (V)] of articulation. Lower panel displays the pooled data. 0 alignment at stop release. VOT values for voiced French stops are mostly artifacts of the labeling procedure as mentioned in the text.

$p < 0.001$). Occlusion duration is 27 ± 9 ms longer in French than in German onsets ($F[1,60] = 8.4$, $p < 0.01$). The interaction of place and language ($F[1,60] = 110.7$, $p < 0.001$) indicates that German onsets show a stronger place effect (25 ± 1 ms; place: $F[1,60] = 534.4$, $p < 0.001$) than French onsets (8 ± 1 ms; $F[1,60] = 34.3$, $p < 0.001$) while on the other hand the effect of language is significant only in velars (35 ± 8 ms; $F[1,60] = 17.1$, $p < 0.001$) but not in labials (18 ± 10 ms; $F[1,60] = 2.7$, $p > 0.05$).There is also an effect of voicing resulting in 4 ± 1 ms longer occlusions in voiceless stops than in voiced ($F[1,60] = 31.1$, $p < 0.001$). In spite of its low variability the effect must be considered rather weak. In fact, as the interaction of voice and place ($F[1,60] = 12.9$, $p < 0.001$) suggests, voicing does not have an effect at all on velars but only on labials (8 ± 1 ms; $F[1,60] = 52.9$, $p < 0.001$). The interaction of voice and language ($F[1,60] = 9.6$, $p < 0.01$) furthermore shows that voicing has no effect on occlusion duration in German but only in French (7 ± 1 ms; $F[1,60] = 28.6$, $p < 0.001$). This leads to the three-way interaction of place, voicing and language ($F[1,60] = 13.4$, $p < 0.001$) which is due to the fact that voicing is only effective on occlusion duration in French labial stops (14 ± 2 ms; $F[1,60] = 79.2$, $p < 0.001$).

The question of how complexity influences occlusion duration is addressed now. Complexity has a main effect causing 9 ± 1 ms longer occlusion in simple than in complex on-sets ($F[1,60] = 121.4$, $p < 0.001$). The interaction of voicing and complexity ($F[1,60] = 6.9$, $p < 0.05$) shows that the voicing effect in French labial stops is further restricted to complex onsets (6 ± 1 ms; $F[1,60] = 21.8$, $p < 0.001$). This is further corroborated by the three-way interaction of place, voicing and complexity ($F[1,60] = 7.8$, $p < 0.01$). The four-way interaction which would round up the picture fails to reach significance by an inch ($F[1,60] = 4.0$, $p = 0.05$).

B. Stop plateau duration

This section lists the effects of place of articulation, language, voicing, and onset complexity thereby addressing expectation E2. A summary of the results is included in Table III. There are main effects of place ($F[1,60] = 88.0$, $p < 0.001$) and voicing ($F[1,60] = 15.6$, $p < 0.001$). The plateau in labials is on average 11 ± 1 ms shorter than in velars and 4 ± 1 ms shorter in voiced than in voiceless stops. Both results can be considered rather weak. This might be connected to the additional main effect of complexity ($F[1,60] = 52.9$, $p < 0.001$) which shortens the plateau duration by about 8 ± 1 ms. Indeed the place effect is stronger in the simple onsets (15 ± 2 ms; $F[1,60] = 61.1$, $p < 0.001$) than in the complex onsets (7 ± 1 ms; $F[1,60] = 28.8$, $p < 0.001$) as the interaction of place and complexity suggests ($F[1,60] = 12.3$, $p < 0.001$) but there are no other interactions involving complexity. Instead there is an interaction of voicing and language ($F[1,60] = 19.1$, $p < 0.001$). Language itself does not have a main effect. However, the voicing effect is restricted to French (10 ± 1 ms; $F[1,60] = 47.5$, $p < 0.001$) and not significant in German. The weak interaction of place and voice ($F[1,60] = 6.1$, $p < 0.05$) points towards a voicing effect in labials only (7 ± 1 ms; $F[1,60] = 82.1$, $p < 0.001$) but not in velars. Finally, an interaction of place and language ($F[1,60] = 5.1$, $p < 0.05$) indicates that the place effect is slightly stronger in French than in German. To summarize, effects on C_1 plateau duration mirror effects on occlusion duration consistently only with regard to complexity. All other factors have different impacts on these two measures.

C. Plateau overlap

This section presents the results concerning expectations E3 and E4 which focus on the effects on plateau overlap. By definition the data under analysis is narrowed to complex onsets. The results are summarized in Table IV. In a first step, the applicability of C_2 plateau duration for normalizing overlap data is investigated.

1. /l/ plateau duration

As mentioned above, C_2 plateau duration was taken into consideration as a possible candidate for normalization of plateau overlap. The reason behind this is the idea that C_2, being the only segmental constant in the consonant clusters considered here, might turn out to be insensitive to variation of C_1 place and voicing as well as the language. It is not.

TABLE IV. Summary of effects C_2 plateau duration and overlap in complex onsets. Interactions are only presented when they contribute crucially to the understanding of the data. Significance codes: *** (p < 0.001); ** (p < 0.01); * (p < 0.05); n.s. (not significant).

	Predictor		
Measure	Place (Vel/Lab)	Voicing (+V/−V)	Language (DE/FR)
C_2 plateau	Vel > Lab **	−V > +V ***	n.s.
Plateau overlap	Vel > Lab ***	+V > −V ***	n.s.
	+V > −V	DE only	

Voicing has a highly significant effect in that /l/ has 5 ± 1 ms longer plateaus after voiceless stops than after voiced ($F[1,60] = 14.5$, $p < 0.001$). Similarly, /l/ plateaus are 4 ± 1 ms longer after velar than after labial stops ($F[1,60] = 10.0$, $p < 0.01$). In spite of their consistency, both effects are obviously rather small. A source of much higher variation is language. While the difference between C_2 plateau durations in German and French is not significant, the languages differ in the strength of variation. The grand means across all speakers of the respective language group and the corresponding standard error are 41 ± 3 ms for French and 57 ± 8 ms for German. The applicability of C_2 plateau duration for overlap normalization is therefore disputable. It is interesting that in the case of C_2 plateau duration, the addition of the random factor for the repetition/block number accounts for a variance of 1.5 ± 1 ms. However, this variance is extremely low compared to the variance explained by the random factor for speakers (169.6 ± 13 ms) and a comparison of models fitted with and without the repetition random factor yields no significant difference (χ^2 [1, $N = 738$] $= 0.72$, $p = 0.3962$).

2. Plateau overlap

A mixed model was designed to calculate the effects of language, voicing and place of articulation on plateau overlap. As outlined in Sec. II C, plateau overlap is the interval between C_1plateau offset and C_2 plateau onset, positive values indicating that the plateaus indeed overlap while negative values indicate a lag between the two plateaus. As with the C_2 plateau duration, the random factor for repetition number does explain some variance (5.0 ± 2.2 ms). However, model comparison again shows that adding this random factor does not improve the fit of the model (χ^2 [1, $N = 738$] $= 3.7$, $p = 0.06$).

Language by itself does not have a significant effect ($F[1,60] = 2.4$, $p > 0.05$). There is a strong main effect of voicing ($F[1,60] = 87.2$, $p < 0.001$) suggesting that there is generally 13 ± 1 ms more overlap in clusters with voiced than with unvoiced stops. The interaction with language, however ($F[1,60] = 37.6$, $p < 0.001$), calls for a closer inspection for each language. For the German data, there is indeed a very significant effect of voicing ($F[1,60] = 100.7$, $p < 0.001$) which accounts for about 21 ± 2 ms more overlap in voiced clusters. The corresponding effect for French is rather marginal ($F[1,60] = 5.9$, $p < 0.05$), the overlap difference between voiced and unvoiced clusters being only 4 ± 2 ms. Examining the data for language specific differences per voicing category shows that clusters with voiceless stops overlap to a similar extent. In clusters with voiced stops, on the other hand, there is a marginally significant difference ($F[1,60] = 6.4$, $p < 0.05$) pointing towards 22 ± 9 ms more overlap in German than in French. Overall, the effect of voicing on overlap is present in German but not in French which confirms the expectations (E3). The interactions of voice and language further provides good evidence that, if anything, overlap in French clusters is more similar to German voiceless clusters but different from German voiced clusters. This is in contrast to expectation E4.

For completeness, we record here the results for place of articulation, though for the reasons given in Sec. II C above they are not a central issue of the present investigation: There is a main effect of place of articulation ($F[1,60] = 34.4$, $p < 0.001$) causing 8 ± 1 ms more overlap in velar + /l/ clusters. There is also an interaction place and voice ($F[1,60] = 7.4$, $p < 0.01$). While the place of articulation effect obtains across both voicing conditions, in clusters with voiced stops the effect is much stronger (12 ± 2 ms: $F[1,60] = 33.2$, $p < 0.001$) than in clusters with voiceless stops (4 ± 2 ms: $F[1,60] = 5.6$, $p < 0.05$).

D. The voiceless phase

Considering the above results, a combination of effects deserves special attention. It seems in Fig. 4 that the combined durations of occlusion and VOT in the case of voiceless stops is rather stable. Figure 5 is a condensed version of Fig. 4 with all voiced tokens removed and an opposition of complex vs simple onsets. Importantly, the onset of the occlusion is used as line-up point in order to better illustrate the relative stability of the voiceless phase. Mainly the timing of the stop's burst within this interval varies as a function of place of articulation and language. For place, of course, this is not a new observation (Weismer, 1980; Cho and Ladefoged, 1999) and it has been argued that underlyingly the glottal devoicing gesture is the same in all cases. Based on findings that show longer VOT in stop + /l/ clusters as compared to simple stop onsets Hoole (2006) discusses several possibilities. The most "radical" possibility proposes lengthening of the glottal gesture due to the addition of the sonorant. To test this here, a mixed model is fitted to a subset of the data including only voiceless stops with place,

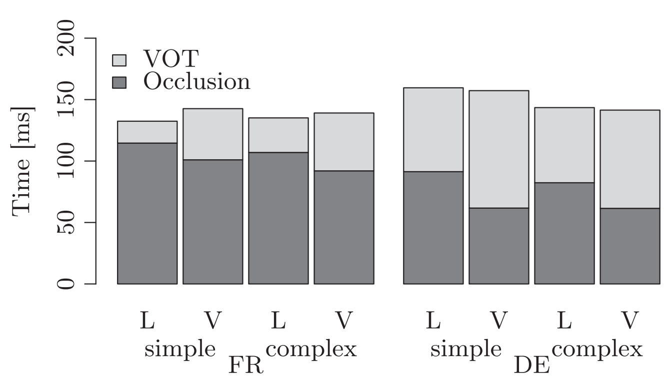

FIG. 5. Mean durations of acoustical occlusion and VOT in simple and complex onsets as a function of language [French (FR) vs German (DE)], complexity (complex vs simple) and place [labial (L) vs velar (V)] of articulation. Alignment at occlusion onset.

language and complexity as predictors and the sum of occlusion and VOT duration as the dependent variable. Place of articulation affects the total voiceless duration such that it is 3 ± 1 ms longer in velar contexts than in labial contexts ($F[1,60] = 5.7$, $p < 0.05$), a weak effect that barely scrapes significance. Language on its own does not affect the duration of the voiceless phase but complexity does ($F[1,60] = 51.9$, $p < 0.001$): Complex onsets have on average 10 ± 1 ms shorter voiceless durations than simple onsets.

The interactions bring language into play. Place and language interact ($F[1,60] = 17.1$, $p < 0.001$) such that the place effect described above is only significant in the French speakers ($F[1,60] = 23.4$, $p < 0.001$) where the voiceless phase is 9 ± 2 ms longer in velar than in labial context. The interaction of language and complexity ($F[1,60] = 25.7$, $p < 0.001$) is due to the fact that the complexity effect above is only significant in the German data where complex onsets have 16 ± 2 ms shorter voiceless phases than simple onsets ($F[1,60] = 71.2$, $p < 0.001$).

IV. SUMMARY AND DISCUSSION

Literature-based expectations concerning VOT and occlusion as outlined in Sec. I B duration were met. The voicing contrast for each language was realized as anticipated: short-lag/long-lag opposition in German vs a voiced/short-lag opposition in French. The data confirm that this patterning also obtained in clusters. It is worth noting that the very short VOT lag after voiced French stops results from the labeling convention applied here: Even in the voiced cases, a VOT interval was labeled should voicing cease during the stop's release. Additionally, one speaker (ff02) regularly produced both /ɡ/ as well as /ɡl/ with a short aspiration phase.

Furthermore, occlusion durations are indeed less in velars than in labials both in single as well as in complex onsets. Reversely, VOT is longer in velars than in labials. The effect of complexity on occlusion duration follows the expectation in that stops in clusters have shorter occlusions than singleton stops.

C_1 plateau durations were expected to show patterns basically similar to those observed for occlusion duration. With regard to complexity the expectations are confirmed: Both occlusions and C_1 plateaus have longer durations in simple than in complex onsets. Voicing also has similar effects on both measures. Occlusion and C_1 plateau are longer in voiceless stops than in voiced stops but this effect is restricted to French labials and—in the case of occlusion duration—to complex onsets. Differences due to C_1 place of articulation result from the different way the active articulators interact with the opposing surface when they form a closure: When the lower lip forms a closure, with the weight of the jaw behind it is likely to carry on moving up after it hits the upper lip (and will only saturate at a plateau if both lips reach their limit of compressibility). On the other hand, the tongue dorsum hits fairly solid structures, which results in a plateau effect in the movement pattern, and thus a tendency to longer plateau durations with the kind of velocity criterion applied here. In any case, as already mentioned, the precise interpretation of differences related to place of articulation is not a central concern of the present paper.

It was assumed that plateau overlap should follow the pattern observed in Hoole *et al.* (2009). Indeed, plateau overlap in German stop + /l/ clusters varies as a function of stop voicing (more overlap/shorter lag in voiced clusters) while it remains stable across both voicing conditions in French. Furthermore, the main question raised in this study was whether overlap in French (both voiced and voiceless) clusters would turn out to be more like voiced or voiceless clusters of German. There is clear evidence in support of the German voiceless pattern, i.e., there is always a considerable lag in French clusters. This is contrary to the argumentation presented in the introduction which was in favor of scenario 1, i.e., overlap in French clusters should pattern as in voiced clusters in German since there is no need to accommodate a glottal gesture/aspiration phase. This surprising result will be further discussed below. It is worth noting here, however, that there does not seem to be a difference of variability as a function of voicing, i.e., neither harmonically voiced clusters (standard error 4.4 ms) nor mixed-voicing clusters (standard error 5.0 ms) exhibit substantially greater stability than the other.

Some further results need to be reviewed that were not explicitly covered by the research questions. Occlusion durations (and along with them the stops' plateau durations) tend to be longer in French than in German. While this was not directly predicted, it is well compatible with the results obtained for VOT and the total phase of voicelessness (in the case of voiceless stops). There are within-language differences between German and French concerning the total voiceless phase (place effect in French, complexity effect in German) but between each other, they do not differ substantially. Concerning occlusion duration and VOT on the other hand the languages differ considerably in such a way that higher VOT and lower occlusion duration in German vs lower VOT and higher occlusion duration in French add up to more or less the same total voiceless duration. In essence this supports previous statements that the timing of the stop release relative to the voiceless phase is fundamentally different between German and French: In French, stop release occurs much later during the voiceless phase than in German. What is new here is that underlyingly French and German stops might have a quantitatively very similar glottal gesture. The results are strongly reminiscent of place-related effects discussed by Hoole (2006) where stop burst occurs earlier in velars than in bilabials within the glottal gesture.

Finally, as seen in Sec. III C, there is a tendency for /l/ plateaus to be shorter in French clusters than in German. The difference is not significant in the mixed model but the languages strongly differ with regard to the extent of variability in /l/ plateau production.

Figure 6 puts the picture together. The sub-figures show the alignment of acoustical (lower bars) and articulatory (upper bars) events separated by place, voicing, complexity and language. The point of departure in this study is clearly visible in terms of overlap relations in the second row of panels for German, where overlap varies as a function of C_1

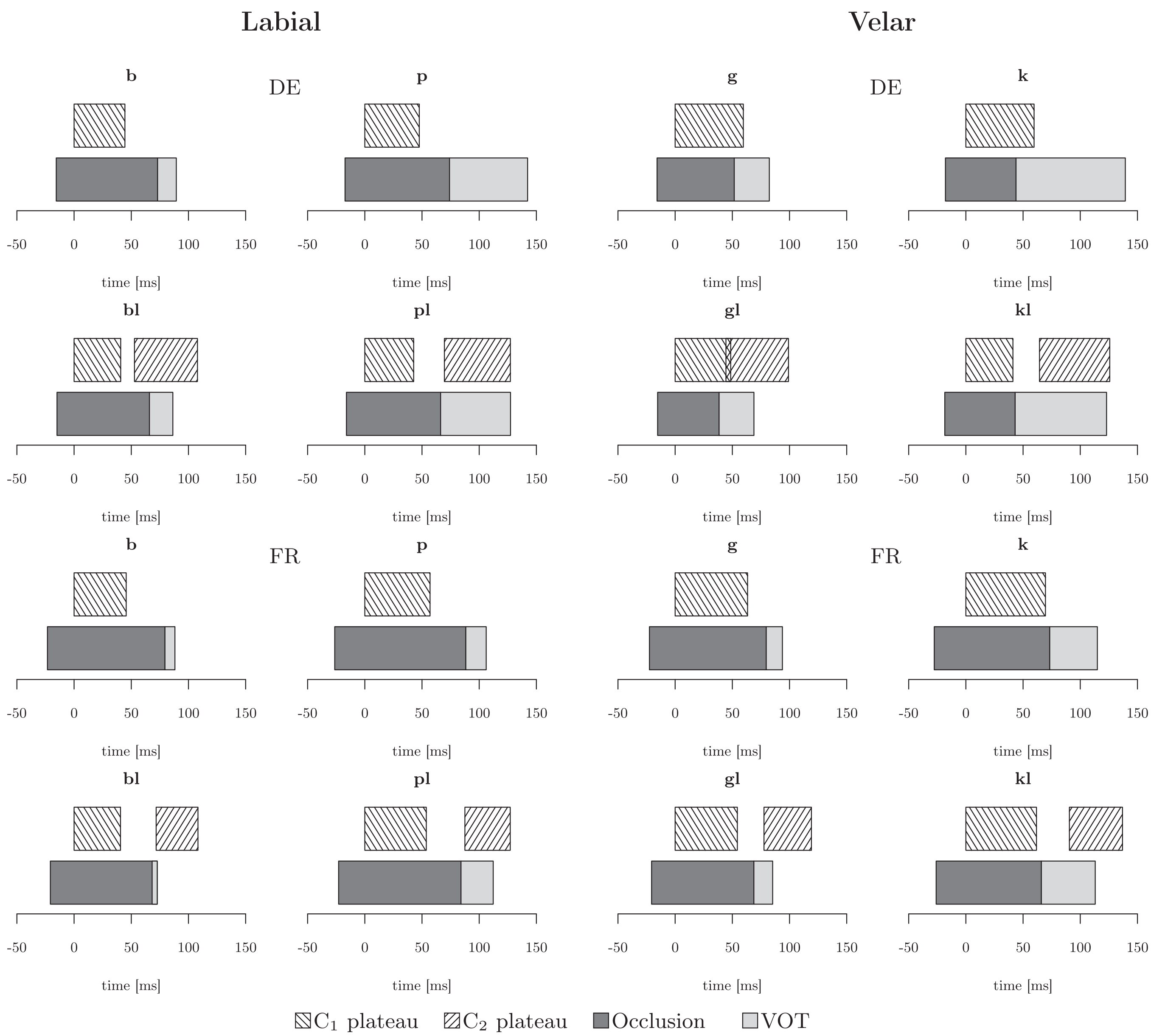

FIG. 6. Occlusion and VOT aligned with articulatory plateaus of singleton stops and stop + /l/ clusters for German (upper four panels) and French (lower four panels). Zero alignment at plateau onset of the stop. Columns on the left show represent labial, on the right the velar data.

voicing, and in the fourth row for French, where overlap is small regardless of voicing. The patterning of acoustic occlusion and VOT in relation to the articulatory landmarks indicates that glottal timing in French clusters is plainly different from the timing in German clusters. This is particularly obvious from the timing of the second consonant. In the Introduction it was argued that C_2 may undergo rightward shift in order to accommodate the glottal gesture. This may or may not be true for German but is evidently not for French where Fig. 6 and the statistics convey the impression that C_2 shifts rightward regardless of the voicing in C_1. In fact the amount of the interconsonantal plateau lag is as large in all French clusters as in the German voiceless clusters in spite of consistently less VOT in French than in German. Furthermore, it appears that the glottal gesture does not extend as far into C_2 in French clusters as it does in German, i.e., French /l/ devoices to a lesser extent. This brings up the question of the domain to which laryngeal properties belong: segments or syllable constituents. In the discussion of German data, Hoole (2006) cites Kehrein and Golston (2004) who conclude their analysis of laryngeal contrast in a large variety of languages with the statement that laryngeal features are properties of the syllable constituents rather than of segments. The German data presented agree with this concept but not the French data where C_2 seems removed from both the stop as well as the devoicing gesture. However, Kehrein and Golston (2004) do not rule out the possibility that within a syllable constituent a laryngeal feature has a stronger association to one segment than to another. In other words, in order to be a property of a syllable onset a laryngeal feature does not necessarily have to spread equally across all segments involved in the onset.

As a final measure to quantify the observed timing differences more precisely, the distance between voice onset and C_2 plateau offset was computed as a percentage of C_2 plateau duration, i.e., the portion of the C_2 plateau that is not devoiced. The data for this measure are restricted to contain only complex onsets to account for the circumstance that simple onsets do not have a C_2 and to mixed voicing clusters since there is no active devoicing gesture in harmonic clusters. Values between 0 and 100 % indicate a voice onset within the constriction plateau of /l/. Values above 100% arise when voicing sets in before C_2 target attainment, negative values when voicelessness outlasts C_2 plateau offset. A mixed model was fitted to this measure as a function of place and language. A corresponding illustration is given in Fig. 7. Significant simple main effects emerge for both predictors, no interactions are encountered. The voiced portion of the C_2 plateau is on average 33 ± 11% shorter after labial than after velar stops ($F[1,60] = 8.8$, $p < 0.01$). More important here, however, is the effect language has on this measure. The C_2 plateau has 78 ± 29% more voicing in French clusters than in German ($F[1,60] = 7.2$, $p < 0.01$). Since in French VOT is in general less and the lag between the consonantal plateaus generally higher than in German, this should come as no surprise. But the voiced portion is bigger in French than in German in spite of the tendency for C_2 duration to be longer in German than in French. In the case of German this result provides evidence for the assumption outlined in the Introduction, that a more continuous sonority modulation might be preferable in the acoustic output. However, one reviewer suggested that the difference in timing between French and German could also be related to French stops being distinctly released (e.g., Fischer-Jørgensen, 1972). This line of thought could be further developed into a perceptual account of cluster timing in French.

The result just described is a further indication that glottal timing in onset clusters considerably depends on language specific grammar: In German, the glottal gesture could be regarded a property of the entire onset (Hoole, 2006) with only marginal voicing at the right edge of the underlyingly voiced sonorant C_2. In the French clusters analyzed here, on the other hand, the glottal gesture appeared to be already receding before C_2 or in other words: C_2 undergoes little devoicing. Interestingly, first results in an ongoing study indicate that this is not the case for /Cr/ clusters.

In recent versions of articulatory phonology (e.g., Pouplier, 2011, for an overview), gestural coordination within syllables is modeled in terms of coupled oscillators. According

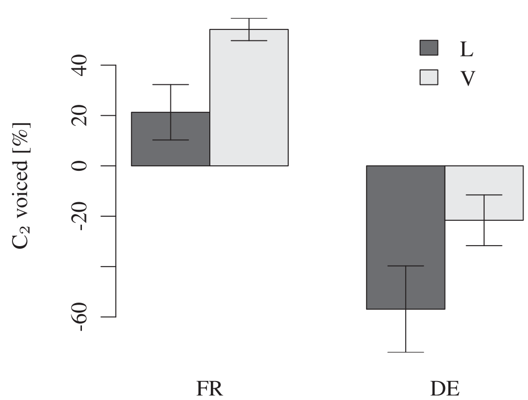

FIG. 7. Voiced portion of the C_2 plateau in complex onsets as a function of language (FR vs DE) and place of articulation (L vs V).

to the model onset-nucleus relations are coupled in-phase, i.e. on the speech planning level they are initiated simultaneously. Nucleus-coda relations and consonant-consonant relations in general are coupled anti-phase, i.e., they are initiated sequentially. For complex onsets, in-phase coupling of all involved onset consonants with the nucleus competes with anti-phase coupling among the onset consonants. On the articulatory surface this competition is reconciled and instead of all involved onset consonants the mid point of the onset (the C-center) emerges in a stable phase relationship with the nucleus. As one reviewer pointed out, this approach can conceivably also account for the present findings on condition that French and German implement different phasing relations of oral onset gestures with laryngeal gestures: While in German all oral onset gestures are uniformly coupled with the glottal gesture, each oral gesture has its own coordination relation with a glottal gesture in French. It would be interesting to explore this idea further with a broader selection of syllable onsets.

Another language other than Georgian (Chitoran *et al.*, 2002) for which an analysis of the interactions of voicing and overlap should be revealing is Moroccan Arabic. It is interesting that French and Moroccan Arabic are similar here since both have fully voiced +V (Zeroual *et al.*, 2006) stops and disprefer overlap in consonant clusters. More generally, one might say that true voicing and overlap are in some way incompatible: /bl/ in German exhibits high overlap but the stop is phonetically not voiced. /bl/ in French is fully voiced but there is very low overlap. A plausible account for this difference can be obtained from aerodynamic considerations. In the case of voiced stop + stop clusters, low overlap would allow for a release of C_1 and thus for a drop of intra-oral pressure. This drop would in turn prevent supra-glottal pressure rising above the threshold where voicing can no longer be maintained for C_2.[2] Word-initially, French does not have stop + stop clusters. However, the existence of such clusters across word or syllable boundaries might have shaped articulatory timing such that low overlap is the pattern generally favored in consonant clusters. This account needs to be substantiated with more data for French. It does find some support, however, from Moroccan Arabic which has voiced stop + stop clusters in initial position and whose basic coordination pattern appears to be low overlap (Gafos *et al.*, 2010; Shaw *et al.*, 2009). A question left open by this account is, as one reviewer pointed out, how coordination patterns across words might be able to affect the typically highly constrained patterns of onset clusters. Another language with a pre-voiced vs short-lag opposition in stop voicing is Slovak. Pouplier and Beňuš (2011) present articulatory data on Slovak showing an overall preference of low overlap in stop + sonorant sequences. However, Pouplier and Beňuš focus on syllabic structure rather than laryngeal specifications and they relate their finding to the frequent occurrence of syllabic sonorants in Slovak. The Georgian data discussed by Chitoran *et al.* (2002) are not contrary to the present account since, as the authors point out, "voiced obstruents in Georgian have very weak voicing, and are not necessarily voiced throughout the closure" (p. 443). In the case of German, these aerodynamic considerations are not necessary since initial underlyingly voiced stops are usually produced without voicing.

In short, although French did not show the overlap pattern initially predicted on the basis of its voicing realization (i.e., that it would pattern like German phonologically voiced stops), this is not because the original explanation originally based on German is wrong, but because for languages with full voicing in phonologically voiced stops an additional constraint comes into play, favoring low overlap. The data presented here show that language specific voicing implementations have an impact on gestural organization and indicate quite clearly that more research is needed to improve the understanding of interactions of laryngeal specification and oral coordination. Also, the inferences made here concerning the timing of the glottal gesture were only based on measurements of acoustical occlusion and VOT. Future research should revisit this issue with techniques which allow for observing both oral and laryngeal articulations.

ACKNOWLEDGMENTS

This work was supported by the DFG Grant No. HO3271/3-1. We also wish to thank Elizabeth Heller, Christine Mooshammer, and Barbara Kühnert.

[1]Speech sounds can be ranked according to the degree of the constrictions involved, where vowels are least constricted, followed by approximants, nasals, fricatives, and finally stops. This ranking has been captured under the term sonority ranking with vowels being most sonorant. Previous research (e.g., Selkirk, 1984) has often found that the world's languages tend to organize syllables such that sonority rises from the margins towards the nucleus (the vowel). This yields an overall preference for syllables such as [plɔŋk] rather than [lpɔkŋ]. However, phonetic research has found no convincing evidence for a physical reality of sonority (e.g., Ohala, 1992).

[2]In addition there are probably cases where strong overlap of C1 and C2 is disadvantageous for maintenance of voicing in C1, e.g., when C2 is dorsal.

Baayen, R. H. (**2008**). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*, 1st ed. (Cambridge University Press, Cambridge, MA), Chap. 7, pp. 241–302.

Benguerel, A. P., Hirose, H., Sawashima, M., and Ushijima, T. (**1978**). "Laryngeal control in French stop production: A fiberscopic, acoustic and electromyographic study," Folia Phoniatr. Logop. **30**, 175–198.

Byrd, D. (**1993**). "54,000 American stops," UCLA Working Papers in Phonetics, URL http://escholarship.org/uc/item/4j43d6qj#page-100 (Last viewed 4/22/13), Vol. 83, pp. 97–116.

Caramazza, A., and Yeni-Komshian, G. H. (**1974**). "Voice onset time in two French dialects," J. Phonetics **2**, 239–245.

Chitoran, I., Goldstein, L. M., and Byrd, D. (**2002**). "Gestural overlap and recoverability: Articulatory evidence from Georgian," in *Laboratory Phonology 7*, edited by N. Warner and C. Gussenhoven (Mouton De Gruyter, Berlin), pp. 419–447.

Cho, T., and Ladefoged, P. (**1999**). "Variation and universals in VOT: evidence from 18 languages," J. Phonetics **27**, 207–229.

Crystal, T. H., and House, A. S. (**1988**). "Segmental durations in connected-speech signals: Current results," J. Acoust. Soc. Am. **83**, 1553–1573.

Docherty, G. J. (**1992**). *The Timing of Voicing in British English Obstruents* (Walter de Gruyter, Berlin), pp. 1–289.

Fischer-Jørgensen, E. (**1972**). " 'p t k' et 'b d g' français en position intervocalique accentuée" ("French 'p t k' and 'b d g' in accented intervocalic position"), in *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, edited by A. Valdman (Mouton, The Hague), pp. 143–200.

Fuchs, S. (**2005**). *Articulatory Correlates of the Voicing Contrast in Alveolar Obstruent Production in German*, Vol. 41 of ZAS Papers in Linguistics (ZAS, Berlin), pp. 1–238.

Gafos, A. I., Hoole, P., Roon, K., and Zeroual, C. (**2010**). "Variation in overlap and phonological grammar in Moroccan Arabic clusters," in *Laboratory Phonology 10*, edited by C. Fougeron, B. Kühnert, M. D'Imperio, and N. Vallée (De Gruyter Mouton, Berlin), pp. 657–698.

Hoole, P. (**2006**). "Experimental studies of laryngeal articulation—Part II: Laryngeal-oral coordination in consonant sequences," Habilitation, Ludwig-Maximilians-Universität, Munich, URL http://www.phonetik.uni-muenchen.de/~hoole/pdf/habilpgg_chap_all. pdf (Last viewed 4/22/13).

Hoole, P., Bombien, L., Kühnert, B., and Mooshammer, C. (**2009**). "Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters," in *Frontiers in Phonetics and Speech Science*, edited by G. Fant, H. Fujisaki, and J. Shen (The Commercial Press, Beijing), pp. 275–286.

Hoole, P., and Zierdt, A. (**2010**). "Five-dimensional articulography," in *Speech Motor Control: New Developments in Basic and Applied Research*, edited by B. Maassen and P. Lieshout (Oxford University Press, Oxford), pp. 331–349.

Kehrein, W., and Golston, C. (**2004**). "A prosodic theory of contrast," Phonology **21**, 1–33.

Klatt, D. H. (**1975**). "Voice onset time, frication, and aspiration in word-initial consonant clusters," J. Speech Hear. Res. **18**, 686–706.

Lisker, L., and Abramson, A. S. (**1964**). "A cross-language study of voicing in initial stops: Acoustical measurements," Word **20**, 384–422.

Lisker, L., and Abramson, A. S. (**1967**). "Some effects of context on voice onset time in English stops," Lang. Speech **10**, 1–28.

Maddieson, I. (**1997**). "Phonetic universals," in *The Handbook of Phonetic Sciences*, edited by W. J. Hardcastle, F. E. Gibbon, and J. Laver (Wiley-Blackwell, Oxford), pp. 619–639.

Maddieson, I. (**2009**). *Patterns of Sounds* (Cambridge University Press, Cambridge, MA), pp. 1–436.

Mooshammer, C., Hoole, P., and Kühnert, B. (**1995**). "On loops," J. Phonetics **23**, 3–21.

Nearey, T., and Rochet, B. (**1994**). "Effects of place of articulation and vowel context on VOT production and perception for French and English stops," J. Int. Phonetic Assoc. **24**, 1–18.

Ohala, J. J. (**1992**). "Alternatives to the sonority hierarchy for explaining the shape of morphemes," in *Papers from the Parasession on the Syllable* (Chicago Linguistic Society, Chicago), pp. 319–383.

Ohala, J. J., and Kawasaki-Fukumori, H. (**1997**). "Alternatives to the sonority hierarchy for explaining segmental sequential constraints," in *Language and its Ecology: Essays in Memory of Einar Haugen*, edited by S. Eliasson and E. H. Jahr (Mouton de Gruyter, Berlin), pp. 343–365.

Pouplier, M. (**2011**). "The atoms of phonological representations," in *The Blackwell Companion to Phonology* (Wiley-Blackwell, Oxford), pp. 107–129.

Pouplier, M., and Beňuš, Š. (**2011**). "On the phonetic status of syllabic consonants: Evidence from Slovak," Laboratory Phonol. **2**, 243–273.

Reubold, U., Harrington, J., and Kleber, F. (**2010**). "Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers," Speech Commun. **52**, 638–651.

Selkirk, E. (**1984**). "On the major class features and syllable theory," in *Language Sound Structure*, edited by M. Aronoff, R. T. Oehrle, F. Kelley, and B. Wilker Stephens (MIT Press, Cambridge, MA), pp. 107–136.

Shaw, J. A., Gafos, A. I., Hoole, P., and Zeroual, C. (**2009**). "Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability in articulation," Phonology **26**, 187–215.

Stevens, K. N. (**2000**). *Acoustic Phonetics* (The MIT Press, Cambridge, MA), pp. 1–617.

Tsuchida, A., and Cohn, A. C. (**2000**). "Sonorant devoicing and the phonetic realization of [spread glottis] in English," working papers of the Cornell Phonetics Laboratory, Vol. 13, pp. 167–181.

Weismer, G. (**1979**). "Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production," J. Phonetics **7**, 197–204.

Weismer, G. (**1980**). "Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations," J. Phonetics **8**, 427–438.

Zeroual, C., Hoole, P., and Fuchs, S. (**2006**). "Etude par transillumination des consonnes occlusives simples et géminées de l'Arabe Marocain" ("Transillumination study of simple and geminate occlusive consonants in Moroccan Arabic"), in *Proceeding of the XXVIth Journées d'Etudes sur la Parole*, pp. 465–468.