Language and Speech
http://las.sagepub.com/

# Temporal Variability and Stability in Infant-Directed Sung Speech: Evidence for Language-specific Patterns

Simone Falk

*Language and Speech* 2011 54: 167 originally published online 31 March 2011
DOI: 10.1177/0023830910397490

*Article*

Language *and* Speech

Language *and* Speech
54(2) 167–180

Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0023830910397490
las.sagepub.com
SAGE

# Temporal Variability and Stability in Infant-Directed Sung Speech: Evidence for Language-specific Patterns

**Simone Falk**
Ludwig-Maximilians-University Munich, Germany

**Abstract**
In this paper, sung speech is used as a methodological tool to explore temporal variability in the timing of word-internal consonants and vowels. It is hypothesized that temporal variability/stability becomes clearer under the varying rhythmical conditions induced by song. This is explored cross-linguistically in German – a language that exhibits a potential vocalic quantity distinction – and the non-quantity languages French and Russian. Songs by non-professional singers, i.e. parents that sang to their infants aged 2 to 13 months in a non-laboratory setting, were recorded and analyzed. Vowel and consonant durations at syllable contacts of trochaic word types with ¦CVCV or ¦CV:CV structure were measured under varying rhythmical conditions. Evidence is provided that in German non-professional singing, the two syllable structures can be differentiated by two distinct temporal variability patterns: vocalic variability (and consonantal stability) was found to be dominant in ¦CV:CV structures whereas consonantal variability (and vocalic stability) was characteristic for ¦CVCV structures. In French and Russian, however, only vocalic variability seemed to apply. Additionally, findings suggest that the different temporal patterns found in German were also supported by the stability pattern at the tonal level. These results point to subtle (supra) segmental timing mechanisms in sung speech that affect temporal targets according to the specific prosodic nature of the language in question.

**Keywords**
German language, non-professional singing, segmental timing, sung speech, temporal targets

## 1 Introduction

In this paper, singing is viewed as a speech register whose prosodic properties should be investigated in relation to spoken speech phenomena. So far, this has not often been done in the (psycho)

**Corresponding author:**
Simone Falk, LMU München, Department für Germanistik, Komparatistik, Nordistik, Deutsch als Fremdsprache, Institut für Deutsche Philologie, Schellingstrasse 3 / RG, D-80799 München, Germany
Email: simone.falk@germanistik.uni-muenchen.de

linguistic, phonetic or (psycho)musicological literature, where song/singing is mostly viewed and researched as professional musical behavior which is claimed to differ fundamentally from verbal behavior in structure and function (e.g., Eggebrecht, 1999). Therefore, a kind of singing in the general population is investigated in this study: singing with infants in the first year of life. At this age parents frequently use both song and speech as communicative tools to address their infants (Trehub & Trainor, 1998). Thus, infant-directed singing can be viewed as a speech register which is not predominantly governed by aesthetical reasons as might be the case in professional operatic singing which makes it interesting for the analysis of segmental structure.

Except for the extensive study of vowel structure and formants (e.g., Sundberg, 1991), we do not find a lot of evidence for the characteristics of segmental structure in song in the current literature. Phonological papers often address the question of how tones and syllables are generally mapped onto each other (e.g., Dell, 2004; Hayes, 2008) and it has been observed that tones (as single pitch events) and syllables in singing are often linked in a one-to-one manner in popular music as, for example, folk or children's songs (Gelber, 1995; Vetterle & Noel, in press), although one-to-many mappings are also frequent in Western classical music. However, these studies leave the question of how the word/syllable-internal timing of vowels and consonants is regulated for further investigation. A relevant parameter might be signal speed. In singing, the whole process of articulation is generally slowed down in comparison with speech. It has been argued that this slow-down is asymmetric in nature affecting especially the vocalic material of the utterance: Kaiser (1983), Eckardt (1999), Stadler Elmer (2002), and Scotto di Carlo (2005) rate the durational predominance of vowels over consonants (5:1 in song, 1:1 in speech according to Eckardt, 1999) as the most significant trait of professionally sung speech. Only Sundberg (1989) alludes to a language-specific timing between vowels and consonants. He suggests that in singing, (phonologically/temporally) long vowels should be followed by short consonants and short vowels by long consonants. This situation is found in spoken Swedish – the native language of Sundberg – and was termed "complementary quantity" (Elert, 1964; Schaeffler, 2005). The phonological system of Swedish features quantity phenomena in both vowels (long vs. short) and consonants (simple vs. geminate). Nevertheless, it is unclear how the timing in sung speech would be regulated in other languages. In this context, a pioneering study by Ross and Lehiste (2001) should be mentioned that used recordings of Estonian runic songs to examine the effects of Estonian speech prosody – and especially prosodic quantity, a three-way length contrast in Estonian – in song. The authors found quantity neutralization at the level of syllables and metric feet, but showed that segmental timing plays an essential role in signaling quantity contrasts in singing. It seems fruitful to look at other languages to gain more insights into segmental timing in sung speech. Therefore, this study will compare sung speech in non-quantity languages (French, Russian) to a language whose quantity status is controversially discussed (German).

The phonological system of German has got a distinctive vowel contrast whose nature is still not well understood. The contrast becomes most obvious in trochaic words of the structure ˈCVCV vs. ˈCVːCV. There are minimal pairs like [ˈraːtən] *raten* '(to) guess' vs. [ˈratən] *Ratten* 'rats' where the meaning of the words is solely distinguished by the duration of the vowel. Phonetically, long vowels in German last twice as long as their short counterparts (Fischer-Jørgensen, 1969). Nevertheless, it has been argued that this duration contrast is not phonologically relevant for all vowels. There are other minimal pairs like [liːt] *Lied* 'song' and [lɪt] *litt* '(he / she / it) suffered' or [ˈfyːlən] *fühlen* '(to) feel' and [ˈfʏlən] *füllen* '(to) fill' where the length contrast is combined with a tense-lax contrast in the vowel. It is unclear if the phonologically primary contrast should be attributed to the qualitative or quantitative aspects of German. Today, many phonologists favor the qualitative contrast because it persists in unstressed syllables whereas the quantity contrast is only found in stressed syllables in German (see Kohler, 1995; Becker, 1998; Hall, 2000).

There is some phonetic evidence adding further aspects to the debate: in a kinematic study, Hoole and Mooshammer (2002) have demonstrated that the acceleration of the articulatory movement shows one maximum in the structures with short/lax vowel whereas there are two maxima (in the transition from the first consonant to the vowel and then from the vowel to the following consonant) in syllables with long/tense vowels. Acoustically, there have long been suggestions in the literature that the energy contour might also be a good candidate to capture the distinctive contrast (Fischer-Jørgensen, 1969; Jørgensen, 1969). Spiekermann (2000) found that the shape of the intensity maximum differs according to syllable structure: in syllables with lax/short vowels it is a peak, whereas in syllables with long/tense vowels it turns out to be a high plateau. However, he assumes that perceptually, length and tenseness of the vowel are the best indicators for the distinction (Spiekermann, 2000, p. 81ff.). The role of durational variability of segments has also been discussed. From a phonological point of view, Restle (2003) makes the assumption that the vowel should be the variable element in closed syllables that have a long/tense vowel, that is, it should be shortened or lengthened depending on speech tempo, whereas the consonant should be the variable element in a closed syllable with a lax/short vowel. In an articulatory study, Hoole, Mooshammer, and Tillmann (1994) found that the vowel is indeed the variable element in CV:C-syllables and always short in duration in CVC-syllables. Nevertheless, in CVC no variability of the consonant could be reliably demonstrated. Fischer-Jørgensen (1969) found very individual patterns of consonant duration in different participants and concluded that it could not be a relevant feature in distinguishing the phonological contrast in German.

In this paper, it is hypothesized that sung speech will highlight the segmental variability pattern of German due to the fact that the rhythmical and tempo structure of song allows for more variable duration conditions than spoken speech. Especially, the role of consonantal variability in the expression of the German phonological contrast will be observed as studies with spoken speech could not demonstrate a reliable influence. German will be compared to French and Russian, two languages without quantity phenomena. In these languages, it is expected that consonants will vary durationally only for expressive purposes. As Sundberg (2000) points out, in sung speech onset consonants can be lengthened in order to highlight the following vowel. This "emphasis by delayed arrival" (Sundberg, 2000, p. 106) is also known in instrumental music where the preceding tonal event is sometimes lengthened or played more forcefully (e.g., on the piano) in order to emphasize the subsequent tone (Sloboda, 1983; Fraňek, 2002). In the following, the timing of sung vowels and consonants in German intervocalic intervals in trochaic words of the structure ¦(C)VC(V) vs. ¦(C)V:C(V) compared to Russian/French ¦(C)VC(V)-intervals will be observed under differing temporal conditions.

# 2 Method and material

## 2.1 Participants

Parents with children aged 2 to 13 months were recruited for a broader fieldwork study on infant-directed singing (Falk, 2009). There were 13 German and 14 Russian speaking parents that participated in this part of the study and 17 French speaking parents. All of them tended to sing quite often with their infants, although none of them was a professional singer or instrumentalist. Recordings were made at home (DAT-Recorder + microphone) while parents sang and spoke to their infants. Parents were asked to interact with their children as they were used to doing in daily life. No further guidelines were given by the researcher. Infants all showed normal hearing capacities and normal progress in cognitive development.

## 2.2 Data

Overall 928 intervocalic intervals in trochaic ¦CV(:)CV structures were extracted from 179 songs sung by the parents in the presence of their infants. The songs, lullabies and play songs, considerably varied in tempo structure. Intervocalic intervals with affricates in the C-position were not included in the analysis because of their controversial status (filling either one or two syllabic positions). There were overall 300 sung performances of 86 different words with ¦CVCV structure in the German data and 211 sung performances of 81 words with ¦CV(:)CV structure. In the French data, there were 81 words sung 203 times by the parents and 108 Russian words sung 214 times.

## 2.3 Analytical method

¦CV(:)CV structures were segmented into phones using Praat (Boersma, 2001). Durations of vowels and consonants at syllable contact were measured and expressed in percentages in relation to the total duration of the VC-interval. Intervocalic intervals were aggregated in each language and averaged over singers according to word context (defined as "same phoneme sequence") and musical context (defined as "same position in the musical phrase, same interval structure, same rhythmical value"). This procedure was used in order to avoid bias because of repetitions of specific VC-intervals in stanzas or songs sung by different singers. As song and text selection was not experimentally controlled, the context and sound environment of the VC-intervals was intensively studied. Vowel height and phrase position as well as consonantal class (manner of articulation) were considered separately. Additionally, in the German material, the location and duration of the stable tonal phase was assessed wherever possible.

# 3 Results

Scatter plots (Figures 1 and 2) show the overall results for durational variation of vowels and consonants in VC-intervals cross-linguistically (raw data).

In German songs (Figure 1), we find an overall weak negative correlation between the duration of the vowel and the consonant (Pearson coefficient: –0.287). This correlation points to a kind of compensatory timing between the two segments in intervocalic intervals during singing: the longer the vowel, the shorter the consonant and vice versa. Figure 1 reveals that in VC-interval structures it is the consonant that shows the greatest durational variability under different rhythmical conditions whereas vowel duration is relatively stable and does not exceed 300 ms (disregarding the few exceptions which will be discussed later). In this condition, the consonant spans the greatest part of the VC-interval. The opposite is true for V:C-interval structures: here, it is the vowel that is durationally variable and consonant duration seems limited to 300 ms. The vowel is generally longer than the consonant.

However, there are a few exceptions to the presented trend: in 15 VC-intervals, the vowel is longer than the consonant. Five items can be traced back to a mother who once had professional voice training (but stopped). Singers are taught to lengthen the vowel to reach an optimal voice volume which decreases vowel intelligibility (e.g., Westerman & Scherer, 2006). This could be the reason for a high vowel proportion in these cases. Five cases were VC-intervals with velar nasals following the low vowel /a/. In those words, a confusion with a similar V:C-word is excluded

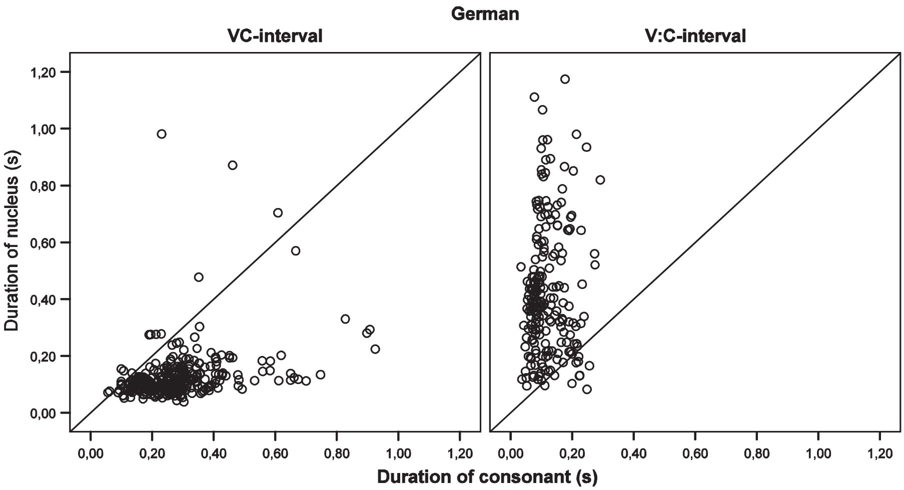


**Figure 1.** German: Duration of vowel and consonant in seconds in VC- and V:C-intervals. Dots on the line: vowel duration equals consonant duration

because the velar nasal only appears after short/lax vowels in German. Thus, recognition of the words is assured independently of the vowel–consonant ratio. The last five cases were the words *alle* 'all' and *tralala* (onomatopoeia) which are characterized by the vowel /a/ followed by a liquid. A similar argumentation to the one given for the velar nasal might hold for these words as well. In V:C-intervals, there are 10 items running counter to the overall tendency of a high vowel proportion in the interval. All these cases contain a voiceless obstruent (see results below), additionally eight of them were produced by a mother that spoke standard High German, but grew up with a Swabian dialect (in Augsburg, Bavaria-Swabia) where the phonological contrast – as in many other Upper German dialects – is phonetically less strictly realized than for example in the northern part of Germany (Spiekermann, 2000).

No correlation was found between vowels and consonants in French and Russian VC-intervals respectively. The results for these languages resemble the German V:C-intervals. Figure 2 shows a similar timing, that is, the vowel is the most variable element in duration. In most of the intervals, the vowel is longer than the consonant; however, there are several cases (41 of 203 in French; 35 of 214 in Russian) where the consonant temporally dominates the VC-interval.

These cases occurred in Russian VC-intervals with voiceless obstruents (especially plosives) and in French VC-intervals with voiceless obstruents and with nasals. Concerning obstruents, it is known that voicing influences the duration of a preceding vowel (House & Fairbanks, 1953; Chen, 1970). The vowel is especially short before voiceless obstruents, which seems to hold for some of the sung data as well (see Figure 5). However, there is no explanation for the variation before nasals in the French material. Noticeably, in French singing, the timing of vowel and consonant in a VC-interval seems to be random sometimes. This is instantiated in Figures 3 and 4 where the singer repeats a stanza in which the word *brune* 'dark' (fem.) occurs in the same musical context. The first time she sings the word, the vowel takes less temporal space of the tonal interval than the consonant (Figure 3); the second time, the proportions of the segments are virtually reversed (Figure 4).

The question of which segment dominates the sung tonal/intervocalic interval was addressed by determining the proportions of vowel and consonant durations in a given V(:)C-interval. Thereafter,

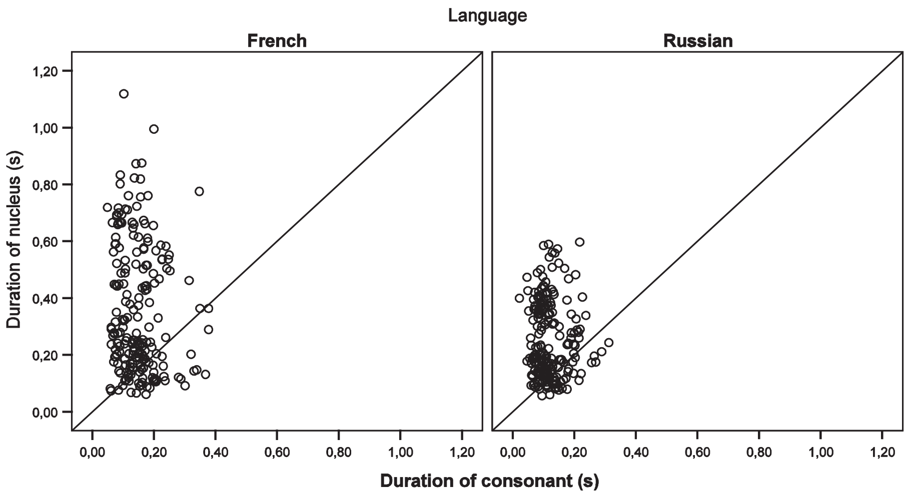


**Figure 2.** French and Russian: Duration of vowel and consonant in seconds in VC-intervals. Dots on the line: vowel duration equals consonant duration

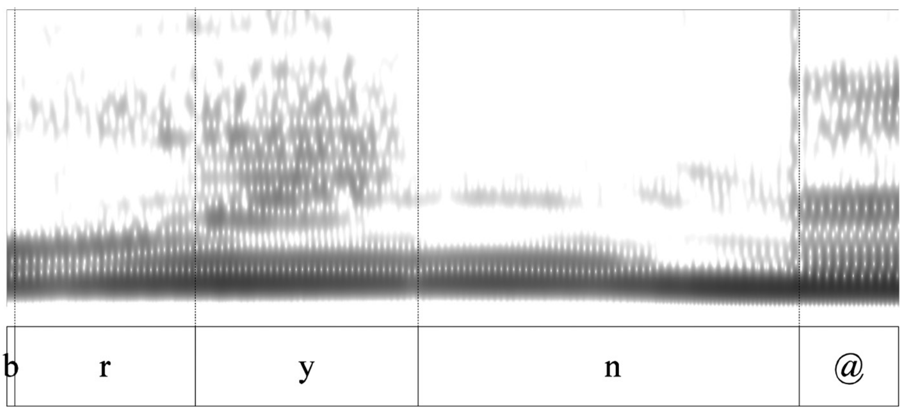


**Figure 3.** Variable vowel duration in French, same word and musical context. Segmented sonogram of the word *brune* 'dark' (generated in Praat, SAMPA-transcription). Vowel proportion: 37% of the VC-interval [yn]

these intervals were aggregated and averaged over singers reducing their number to 128 VC- and 132 V:C-intervals in German, to 104 intervals in French, and to 138 intervals in Russian. Aggregation followed the guidelines described in the 'Analytical method' section. In the aggregated French and Russian material, items were grouped in order to assess which contextual factors could bias the proportion of the vocalic part of sung VC-intervals. Three factors were to be taken into account. First, it is widely known in the literature that vowels as well as musical tones undergo lengthening in phrase-final position (e.g., Oller, 1973; Lindbløm, 1978; Vaissière, 1991). In this study, this parameter was especially important for French as many intervals occurred at the end of phrases. This is due to the fact that words with trochaic structure are unusual in spoken French which has got a final accent. In poetic and sung discourse, however, final unstressed Schwa syllables can or even must be articulated

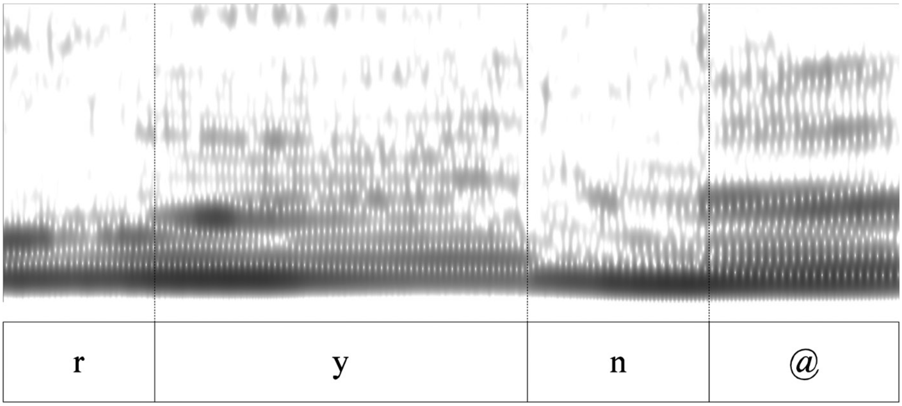


**Figure 4.** Variable vowel duration in French, same word and musical context. Segmented sonogram of the word *brune* 'dark' (generated in Praat SAMPA-transcription). Vowel proportion: 67% of the VC-interval [yn]

(Dell, 1989). This is frequently the case at the end of phrases. Second, low vowels are inherently longer than high vowels (Elert, 1964; Lehiste, 1970). Third, as mentioned above, voicing of the following consonant has an effect on vowel duration in spoken speech (House & Fairbanks, 1953). Thus, the items in French and Russian were analyzed according to the influence of these parameters on the vowel proportion in the VC-interval: interval position in the phrase (phrase-final, non-phrase-final), vowel height (high, middle, low), voicing of consonant (voiced, voiceless).

A Mann-Whitney U test was performed in French and Russian for the parameters "Interval position in the phrase" and "Voicing of consonant" and a Kruskal-Wallis test for "Vowel height". In Russian, phrase-finality had an effect on the proportion of the vowel, $p < 0.05$, $U = 1459$, $n_{1(final)} = 41$, $n_{2(non\text{-}final)} = 97$, thus the phrase-final items were excluded in the Mann-Whitney U test on "Voicing of consonant" and the Kruskal-Wallis test on "Vowel height". Additionally, four Russian items containing the vowel [ɨ] were excluded from analysis. Only "Voicing of consonant" had a significant effect on the vowel proportion in the VC-interval in both languages (French: $p < 0.001$, $U = 512$, $n_{1(voiceless)} = 35$, $n_{2(voiced)} = 69$; Russian: $p < 0.001$, $U = 423$, $n_{1(voiceless)} = 34$, $n_{2(voiced)} = 59$).

Results for the aggregated data were plotted in regard to consonantal class (obstruents: voiced, voiceless; sonorants). Figure 5 gives the results for the consonantal proportion in VC-intervals in French and Russian. As can be seen, variance is very high in all three consonantal classes, though the median of consonantal proportions stays overall under 50%. The median for voiceless obstruents is highest. Voiceless obstruents, especially plosives, can reach up to 77% of the tonal interval in French and up to 69% in Russian.

German aggregated VC- and V:C-intervals have been compared to these results. Both interval groups were analyzed with respect to the aforementioned parameters. Neither the parameter "Position in the phrase" nor "Voicing of consonant" differed in the interval groups. A bias by these parameters could be excluded. Concerning "Voicing of consonant", the analysis revealed a preponderance of voiced/sonorant consonants in V:C-intervals. The difference between the groups is most obvious when obstruents are considered. As Table 1 shows, an almost complementary distribution of voiced/voiceless obstruents exists (a correlation which can be explained from a historical point of view, see Becker, 2008).

It is all the more surprising that the difference between VC- and V:C-intervals in German remains visible across consonantal class (see Figure 6). In German VC-intervals the

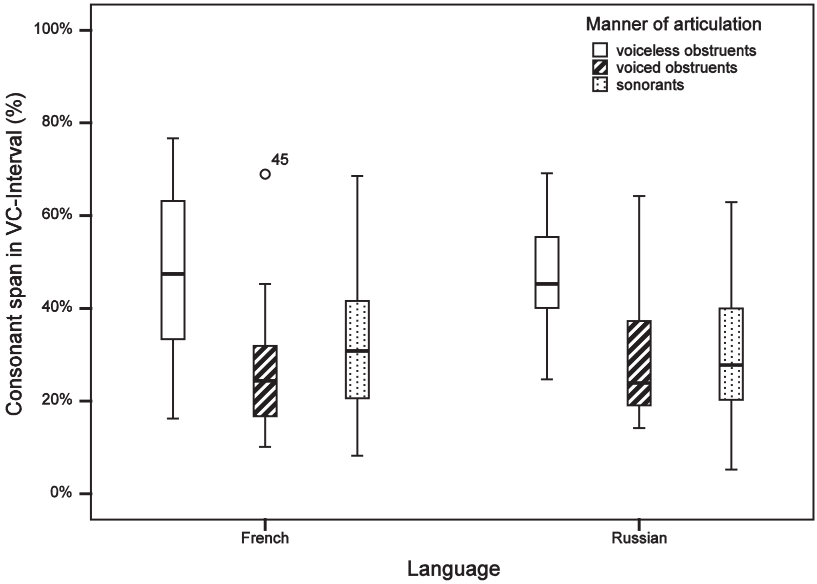


**Figure 5.** Boxplots of consonant proportions in tonal intervals (VC-interval) with respect to consonantal class. Results for French and Russian. Aggregated data

**Table 1.** Distribution of voiced/voiceless obstruents in German VC- and V:C-intervals. Aggregated data

| Interval | Voicing | Labial plosive | Alveolar plosive | Velar plosive | Fricative | **Total** |
|---|---|---|---|---|---|---|
| VC | voiceless | 4 | 14 | 18 | 22 | **58** |
| | voiced | 2 | 0 | 0 | 0 | **2** |
| V:C | voiceless | 0 | 8 | 2 | 28 | **38** |
| | voiced | 26 | 18 | 13 | 9 | **66** |

consonant proportion is significantly higher (median in all consonantal classes over 60% of the total interval) than in V:C-intervals where the median is lower than 40% of the interval. Nevertheless, the variance of the proportion of voiceless obstruents in V:C-intervals is high, reaching up to 89%.

These results are supported by statistical analysis: a Mann-Whitney U test on both interval groups is highly significant, overall, $p < 0.001$, $U = 182$, $n_{1(VC)} = 128$, $n_{2(V:C)} = 132$, and for voiceless obstruents, $p < 0.001$, $U = 43$, $n_{1(VC)} = 58$, $n_{2(V:C)} = 38$, and sonorants, $p < 0.001$, $U = 2$, $n_{1(VC)} = 68$, $n_{2(V:C)} = 28$, tested separately (voiced obstruents are not tested because of the small number of items in VC-intervals).

French, Russian and German data are compared in Figure 7. Only results for nasals (without velar nasals) and liquids are presented. There are 50 French, 43 Russian, 55 German VC- and 28 aggregated V:C-intervals. German tonal VC-intervals are clearly dominated by consonantal material. The French and Russian results resemble German V:C-intervals, but the variance of the consonantal portion in French and Russian is considerably higher.

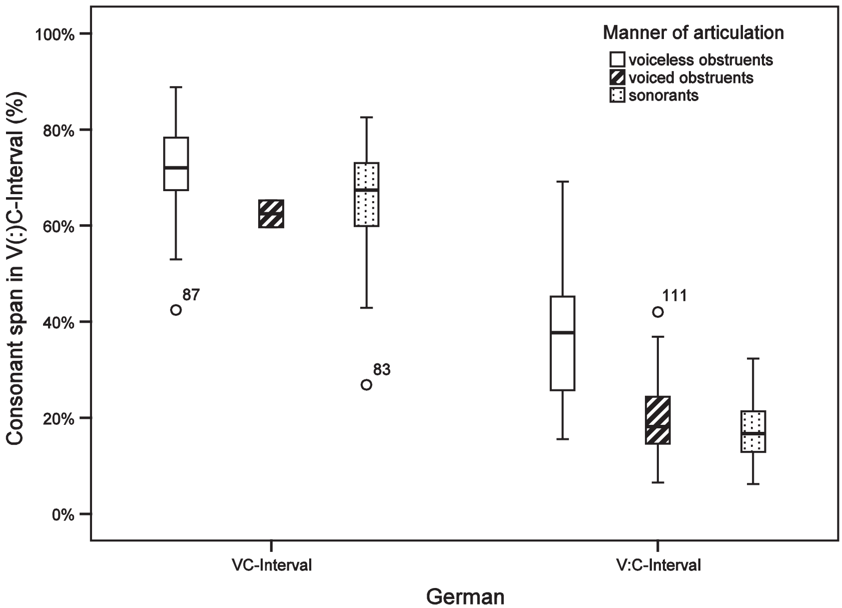


**Figure 6.** Boxplots of consonant proportions in German V(:)C-intervals with respect to consonantal class. The box representing voiced obstruents in VC-intervals may be disregarded because there are only two items in this group. Aggregated data

Finally, the question was addressed of which segment of the German VC-intervals can be tone-bearing in sung speech. Therefore, the duration and location of the stable tonal portion and tonal transitions in the German V(:)C-intervals were measured. First, this had to be done with V(:)C-intervals, where the C was a nasal or liquid that could potentially bear a periodic tone. Second, transitions and stable tonal phases could only be assessed if there was an interval of at least a semitone between two following tones. Results were as follows: in 32 V:C-intervals that met these conditions, the stable tonal phase occurs exclusively in the vowel, the transition to the next tonal value either begins in the last fifth of the vowel continuing through the consonant or it takes place completely in the consonant. By contrast, the picture in VC-intervals is different: in 52 cases out of 55, the consonant carries an important part of the stable tonal phase (at least 20%). More than half of those cases (36) actually include over 50% up to 100% of the stable tonal phase. A sonogram of such a German VC-interval (C: nasal) is presented in Figure 8. The frequency curve shows that it is the consonant that carries most of the stable tonal phase in this interval. In this example, the transition to the next tone takes place in the very last part of the consonant. Thus, it may be concluded that consonants not only shape the rhythmical value of a tone but can also at times carry the crucial tonal information of a note in German infant-directed singing.

## 4 Discussion

Results confirm that German sung VC- and V:C-intervals show distinct patterns of vocalic and consonantal durational variability which can be ascribed to the phonological contrast in German. The most important finding is that in VC-intervals with short/lax vowel, it is indeed the consonant

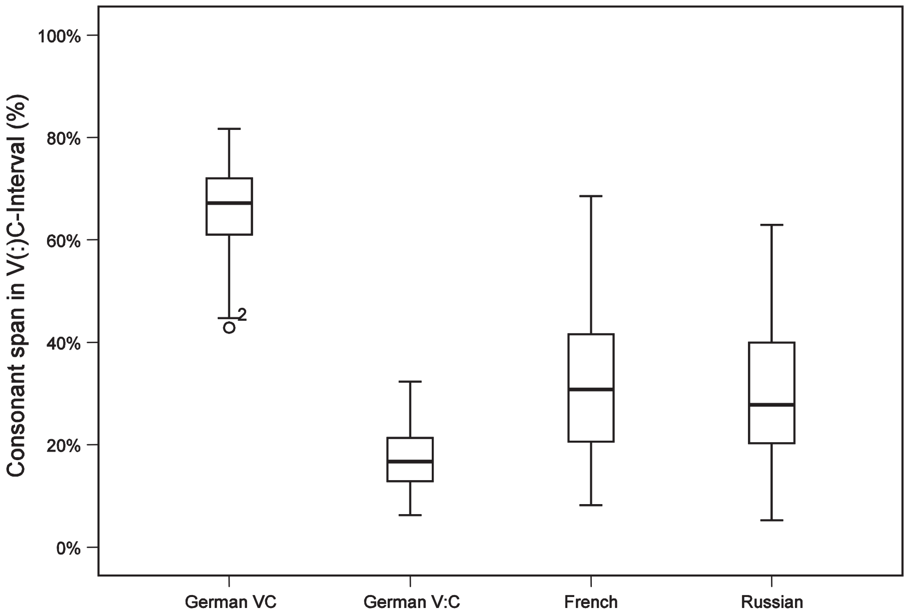


**Figure 7.** Boxplots of consonant proportions in German, French and Russian, nasals and liquids only (without velar nasals). Aggregated data

that is variable in duration under different rhythmical conditions, whereas the vowel is temporally stable. In German V:C-intervals with long/tense vowel, the vowel is by default the variable segment and tone-bearing element and the consonant is the stable part. This pattern of mutual segmental temporal variability and stability could not be demonstrated in previous studies with spoken speech, but becomes apparent in song that expands the temporal range of articulatory activity. These results indicate that temporal targets in German undergo considerable variation in the process of speech, though very distinct variability patterns for VC/V:C-structures emerge under specific circumstances as it was shown in this study with sung speech. The findings are also in line with phonological theorizing that underlines the importance of the dynamic link/coupling between vowel and consonant in indicating the distinctive contrast in German (as it is the case for proponents of the so-called "syllable-cut theory", e.g., Vennemann, 1991; Becker, 1998; Restle, 2003; relying on Sievers, 1901; Jespersen, 1913). It would be worthwhile to clarify which factors ultimately guide perception of the phonological contrast in German sung speech compared to spoken speech. Temporal interaction between vowels and consonants at syllable boundaries might be an important factor, at least in the word groups investigated in this study.

Furthermore, the findings challenge the widely held view in the literature that vowels should always be the most prominent and dominant segments in singing and support the idea of Sundberg (1989) that there can be mutual interaction between the duration of vowels and consonants in sung speech. Segmental timing in singing also seems to depend on the quantitative nature of individual languages. Ross and Lehiste (2001) already provided evidence that quantity contrasts in Estonian songs can be solely regulated at the segmental level, especially by the duration of intersyllabic consonants. Respective temporal variability of vowels and consonants at syllable boundaries was found to be characteristic for German in this study. However, the pattern of German segmental

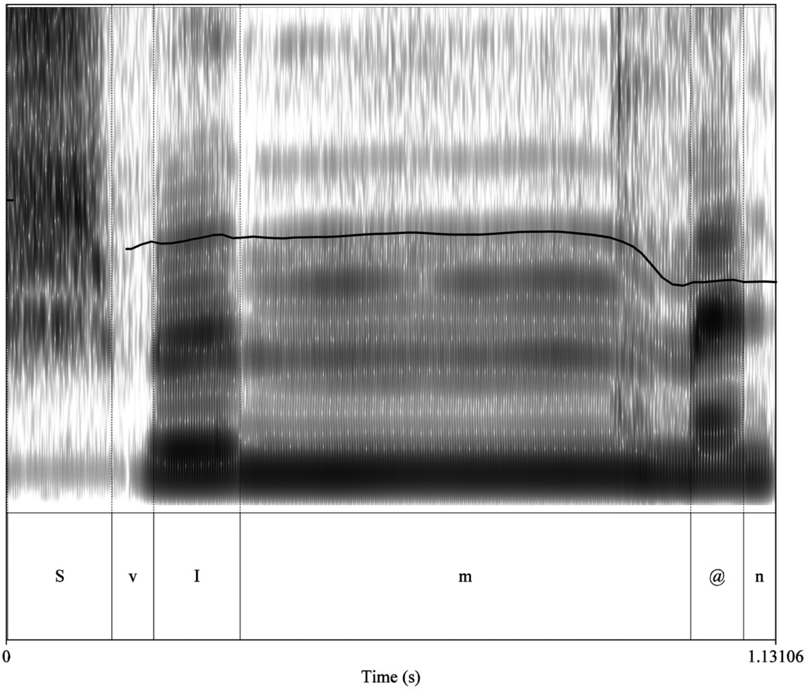


**Figure 8.** Sonogram of the sung word *schwimmen* 'to swim' and frequency curve (270 Hz, overall range: 500 Hz). The nasal is lengthened and bears most of the tonal information of the tonal interval

timing differed clearly from the pattern found in non-quantity languages like French and Russian. Here, it is the vowel that generally dominates the intervocalic interval, even though consonantal segments in these two languages sometimes vary considerably in duration. This might be due to expressive purposes (Sundberg, 2000; Sundberg & Bauer-Huppmann, 2007). It seems plausible that a consonant might be lengthened to achieve a staccato effect or to serve rhythmical accentuation. This could be most pronounced in (play)songs that are rhythmically complex or combined with rhythmical movements during singing. As shown in the Results section, some French parents vary the duration of vocalic and consonantal material at will in VC-intervals that appear in the same song and musical context. Such a deliberate durational variability might not be found in German where duration serves a phonological distinction. In German other factors like intensity or even tonal fluctuations such as jitter and shimmer or other tonal aspects might be used to serve these expressive functions. This must be left to further investigation. It would also be valuable to further investigate the characteristics of German sung speech in comparison to a quantity language with vocalic and/or consonantal length distinctions.

Since the material was taken from a fieldwork study, biasing context factors had to be identified. One important aspect was voicing of consonant and consonantal class. As in spoken speech (House & Fairbanks, 1953; Chen, 1970), vowels in sung speech are shorter before voiceless obstruents, especially before plosives. Position in the phrase might also play a role as shown in the Russian material. It could be refuted that these factors were at the bottom of the durational variation in German V(:)C-intervals. However, it is a lexical fact that German short/lax vowels occur more often with a following voiceless obstruent whereas long/tense vowels are followed more often by voiced obstruents. From a developmental point of view, this segmental correlation could help a child learner to acquire the phonologically relevant contrast more readily. Because of its slower articulation and the VC-timing relations, sung speech could be particularly suitable to render the contrast perceptually salient. Perceptual testing with infants will be necessary to clarify this hypothesis.

Another contextual factor that slightly blurred the picture in German was the occurrence of VC-intervals with velar nasals or some (pseudo)words with the vowel /a/ + following lateral (*alle*, *tralala*). For these words no V:C counterpart exists that would express another meaning. Thus, parents might not be as strict in their VC-timing as in other intervocalic intervals where the meaning could be potentially misidentified. Another factor that should be considered in further studies is dialectal provenance. In this study, all the participants spoke standard High German; however, the phonological contrast between short/lax and long/tense vowels is not equally pronounced in all German regions. This might have been the reason for some exceptions to the general pattern found in one mother that grew up with a dialectal background.

Additionally, this study demonstrated that sonorant consonants can also carry the crucial part of tonal information needed to identify tonal values in German sung speech. Although further measurements and more data will be needed, this observation alludes to subtle timing mechanisms that operate at both the rhythmical and the tonal level in sung speech to express language-specific segmental characteristics. Perceptual studies on this aspect would also be very helpful. The data of this study are borne out by a further study of Falk (2009) with the same fieldwork material. A total of 352 infant-directed sung vowels were examined that are unstable at the tonal level, i.e. vowels that exhibit an unstable tonal "glide" before the stable tonal phase begins. Tonal glides in vowels are reported to be very salient for infant listeners (Trainor & Desjardins, 2002). The beginning of the stable tonal phase after such a glide shows some specificities in the German material: in almost 85% of the syllables with short/lax vowels and following sonorant consonants the stable tonal phase was located in the consonant, whereas in syllables with long/tense vowels the stable tonal phase was never located in consonantal material.

Finally, sung speech appeared to be a suitable tool to uncover the scope of segmental variation. This might have been due to the slower articulation speed and greater rhythmical variability induced by the musical characteristics of sung speech. The advantage of sung speech over slow speech elicited by instructions in an experimental setting might be that it is produced more naturally and thus provides more consistent results – especially with non-professional singers. There is reason to believe that non-professional singers behave differently than professional singers or persons who have had intensive voice training. This could be one way of explaining the discrepancy between previous studies on segmental duration in singing and this investigation. Professional singers could be inclined to give up language-specific timing at the segmental level for greater volume and sonority of their voice. Parents singing to their infants rather seem to preserve timing and textual meaning. Although infants at this age are not yet able to understand the exact meaning of sung texts, parents seem to make a point of producing a coherent and comprehensible text/narrative in sung speech. This becomes especially evident whenever parents substitute unknown words in traditional songs.[1]

Overall, the findings of this study suggest that it will be very valuable in the future to explore the effects of singing on language acquisition. The difference between professional and non-professional singers and their communicative intentions should also be further investigated. To conclude, sung speech might be a good tool in experimental studies to examine the characteristics and realization of temporal targets cross-linguistically and to gain more insights into the nature of segmental timing relations – at the articulatory and perceptual level.

## Acknowledgements

This research has been supported by the Studienstiftung des deutschen Volkes. Thanks for their helpful remarks and comments to Keith Johnson, Fred Cummins, Richard Rhodes and Phil Hoole.

## Note

1 In the German lullaby "Guten Abend, gut Nacht" ('Good evening, good night' by Johannes Brahms) there is a verse line that states that the bed/pillow is decorated with flowers which are described by an old German word for lilac *Näglein*. This term could be misinterpreted as denoting small nails. Most of the parents replace the old word by another word for a flower.

## References

Becker, T. (1998). *Das Vokalsystem der deutschen Standardsprache*. Frankfurt a.M.: Lang.

Becker, T. (2008). Akzent und Vokalwandel seit althochdeutscher Zeit. *Beiträge zur Geschichte der deutschen Sprache und Literatur, 130*(3), 401–419.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341–345.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22*, 129–159.

Dell, F. (1989). Concordances rythmiques entre la musique et les paroles dans le chant. L'accent de l'*e muet* dans la chanson française. In M. Dominicy (Ed.), *Le souci des apparences* (pp. 121–136). Bruxelles: Editions de l'Université de Bruxelles.

Dell, F. (2004). Singing as counting syllables: Text-to-tune alignment in traditional French songs. In *Proceedings of the Conference Words and Music, University of Missouri-Columbia, March 14, 2003*. Missouri, Columbia.

Eckardt, F. (1999). *Singen und Sprechen im Vergleich artikulatorischer Bewegungen*. Darmstadt: THIASOS Musikverlag.

Eggebrecht, H. H. (1999). Musik und Sprache. In A. Riethmüller (Ed.), *Sprache und Musik. Perspektiven einer Beziehung* (pp. 9–14). Laaber: Laaber.

Elert, C. C. (1964). *Phonologic studies of quantity in Swedish*. Uppsala: Almqvist & Wiksell.

Falk, S. (2009). *Musik und Sprachprosodie. Kindgerichtetes Singen im frühen Spracherwerb*. Berlin: de Gruyter.

Fischer-Jørgensen, E. (1969). Untersuchungen zum sogenannten festen und losen Anschluss. In K. Hyldgaard-Jensen & S. Steffensen (Eds.), *Kopenhagener germanistische Studien*, Vol. 1 (pp. 138–164). Kopenhagen: Akademisk Forlag.

Fraňek, M. (2002). *Rhythm production and perception: The role of accentuation and tempo*. Prag: Editio St Aegidii.

Gelber, L. (1995). *Le chant enfantin. Étude psycho-génétique*. Louvain-la-Neuve: Département d'archéologie et d'histoire de l'art, collège Érasme.

Hall, A. T. (2000). *Phonologie. Eine Einführung*. Berlin: de Gruyter.

Hayes, B. (2008). Faithfulness and componentiality in metrics. In K. Hanson & S. Inkelas (Eds.), *The nature of the word: Essays in honor of Paul Kiparsky* (pp. 113–148). Cambridge, MA: MIT Press.

Hoole, P., & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In P. Auer, P. Gilles, & H. Spiekermann (Eds.), *Silbenschnitt und Tonakzente* (pp. 129–152). Tübingen: Niemeyer.

Hoole, P., Mooshammer, C., & Tillmann, H. G. (1994). Kinematic analysis of vowel production in German. In *Proceedings of the ICSLP 94* (pp. 53–56). Yokohama: ISCA.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America, 25*(1), 105–113.

Jespersen, O. (1913). *Lehrbuch der Phonetik*. Berlin: Verlag von B.G. Teubner.

Jørgensen, H. P. (1969). Über den Intensitätsverlauf beim sogenannten festen und losen Anschluss im Deutschen. In K. Hyldgaard-Jensen & S. Steffensen (Eds.), *Kopenhagener germanistische Studien*, Vol. 1 (pp. 165–186). Kopenhagen: Akademisk Forlag.

Kaiser, J. F. (1983). Some observations on vocal tract operation from a fluid flow point of view. In I. R. Titze & R. C. Scherer (Eds.), *Vocal fold physiology: Biomechanics, acoustics and phonatory control* (pp. 358–386). Denver, CO: The Denver Center for the Performing Arts.
Kohler, K. J. (1995). *Einführung in die Phonetik des Deutschen* (2nd ed.). Berlin: ESV.
Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
Lindbløm, B. (1978). Final lengthening in speech and music. In E. Garding, G. Bruce, & R. Bannert (Eds.), *Nordic prosody: Papers from a symposium* (pp. 85–101). Lund: Lund University.
Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America, 54*(5), 1235–1247.
Restle, D. (2003). *Silbenschnitt – Quantität – Kopplung. Zur Geschichte, Charakterisierung und Typologie der Anschlußprosodie*. München: Fink Verlag.
Ross, J., & Lehiste, I. (2001). *The temporal structure of Estonian runic songs*. Berlin: de Gruyter.
Schaeffler, F. (2005). *Phonological quantity in Swedish dialects: Typological aspects, phonetic variation and diachronic change*. Dissertation, Umea University.
Scotto di Carlo, N. (2005). Contraintes de production et intelligibilité de la voix chantée. *Travaux Interdisciplinaires du Laboratoire Paroles et Langage, 24*, 159–179.
Sievers, E. (1901). *Grundzüge der Phonetik*. Leipzig: Härtel & Breitkopf.
Sloboda, J. A. (1983). The communication of musical metre in piano performances. *Quarterly Journal of Experimental Psychology, 35A*, 377–396.
Spiekermann, H. (2000). *Silbenschnitt in deutschen Dialekten*. Tübingen: Niemeyer.
Stadler Elmer, S. (2002). *Kinder singen Lieder: Über den Prozess der Kultivierung des vokalen Ausdrucks*. Münster: Waxmann.
Sundberg, J. (1989). Synthesis of singing by rule. In M. V. Mathews & J. R. Pierce (Eds.), *Current directions in computer music research* (pp. 45–55). Cambridge, MA: MIT Press.
Sundberg, J. (1991). The acoustics of the singing voice. In W. Wang (Ed.), *The emergence of language: Development and evolution* (pp. 104–116). New York: Freeman and Company.
Sundberg, J. (2000). Emotive transforms. *Phonetica, 57*, 95–112.
Sundberg, J., & Bauer-Huppmann, J. (2007). When does a sung tone start? *Journal of Voice, 21*, 285–293.
Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review, 9*(2), 335–340.
Trehub, S. E., & Trainor, L. (1998). Singing to infants: Lullabies and play songs. *Advances in Infancy Research, 12*, 43–78.
Vaissière, J. (1991). Rhythm, accentuation and final lengthening in French. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, language, speech and brain: Proceedings of an international symposium at the Wenner-Gren Center Stockholm, 5–8 September 1990* (pp. 108–120). Basingstoke: Macmillan.
Vennemann, T. (1991). Skizze der deutschen Wortprosodie. *Zeitschrift für Sprachwissenschaft, 10*(1), 86–111.
Vetterle, R., & Noel, P. A. H. (in press). Bavarian *Zwiefache*: Investigating the interface between rhythm, metrics and song. In J. L. Aroui and A. Arleo (Eds.), *Towards a Typology of Poetic Forms*. Amsterdam: Elsevier.
Westerman, J. G., & Scherer, R. C. (2006). Vowel intelligibility in classical singing. *Journal of Voice, 20*(2), 198–210.