Accentuation cues in French and German

Jacques Koreman[2], Bistra Andreeva[1], William Barry[1]

[1] Saarland University, Saarbrücken, Germany
[2] Norwegian University of Science and Technology (NTNU), Trondheim, Norway
jacques.koreman@hf.ntnu.no; andreeva/wbarry@coli.uni-saarland.de

Abstract

The use of acoustic properties by French and German speakers in accentuating and de-accentuating words in accordance with changes in information weighting is examined. In particular, the role of duration is shown to differ for the two languages, confirming earlier findings. In addition, a difference in the relative contribution of F_0, intensity and spectral balance is found for the prediction of accentuation in the two languages. The link with phonological differences between the languages and questions of analysis methodology are discussed.

1. Introduction

French and German belong to different branches of the Indo-European family of languages and are generally considered to have different rhythmic structures, However, in order to vary the prominence of the syllables and words in an utterance and create its intended prosodic pattern, both languages are restricted to modifying the same acoustic properties of the speech signal – namely duration, F_0, intensity and spectral definition. In many studies [e.g., 2, 4, 6], the pre-eminent contribution of F_0 and duration to the variation in the level of accentuation is reported, with intensity being accorded a very minor role (but see [5]). Some languages exploit a phonologically established vowel reduction strategy to signal unstressed syllables at the lexical level [3], but such reduction processes are not considered to support phrase-level de-accentuation of lexically stressed syllables.

Parallel to the production of different levels of prominence that serve the prosodic marking of information structure (e.g. topic accent, focal accent, contrastive or corrective accent, etc.), languages have to conform to the regularities of their segmental phonologies. Thus, in languages with free lexical stress, the lexically stressed syllable of a non-monosyllabic word has to be marked and this needs to be kept separate from the phrase-level accent. If tonal patterns are used to signal lexical stressing (as e.g. in Norwegian, Swedish, Japanese, etc.), there may be some constraints on the use of F_0 for phrase-level accentuation. Similarly, if a language has segmental length oppositions, such as vowel length in Czech, Finnish, German, etc., or geminate consonants, as in Italian and Finnish or Norwegian and Swedish (where there is an interdependence with the length of the preceding vowel), the exploitation of duration for accentual marking may be constrained.

The assumption of different constraints across languages lay behind [1], in which the degree of change in duration, F_0, intensity and spectral definition as phonetic exponents of different levels of accentuation was compared in French and German. It was shown that, while French made comparable use of F_0 and duration in differentiating between de-accented and accented words, German exploited duration (i.e., shortening and lengthening) a lot less strongly. This result was seen as confirmation of a constraint resulting from the use in German of a vowel-length distinction at the segmental level.

The aim of this paper is to consider in more detail the contribution of the individual parameters to accentual differentiation in the two languages. The statistical weighting of the parameters in their contribution to the differentiation of a word as "accented" or "de-accented" is of particular interest.

2. Speech material and data analysis

2.1. Speech material

As a basis for the direct comparison of parameter values across different levels of phrasal accentuation, utterances with de-accented and accented variants of the *same words* were used. Short sentences were constructed containing two one- or two-syllable "critical words" (CWs), one early (but not initial) and one late (but not final) in the sentence. The sentences for German (with the CWs underlined) are:

1. Der <u>Mann</u> fuhr den <u>Wagen</u> vor.
2. Das <u>Bild</u> soll nicht <u>hässlich</u> sein.
3. Das <u>Kind</u> sollte im <u>Bett</u> sein.
4. Der <u>Peter</u> kann den <u>Film</u> gucken.
5. Das <u>Mädchen</u> soll ein <u>Bild</u> malen.
6. Mein <u>Vater</u> kann <u>Türkisch</u> lesen.

The French sentences are:

1. Mon <u>fils</u> met les <u>vases</u> par terre.
2. Mon <u>neveu</u> fait du <u>thé</u> le matin.
3. Les <u>parcs</u> sont <u>fermés</u> ce soir.
4. Tes <u>pommes</u> sont <u>belles</u> ce matin.
5. Le <u>dentiste</u> s'est <u>cassé</u> la jambe.
6. Mon <u>mari</u> mange du <u>pain</u> ce soir.

For each sentence, questions were devised to elicit a response with a *non-contrastive narrow-focus* a) on the early and b) on the late CW.

To provide a basis for comparing the parameter modification across sentences and languages independently of the differing segmental structuring of the critical words, a reiterant "*dada*" version of each realization was produced immediately after the normal-text response.

2.2. Speakers, recordings and analysis

Six speakers of northern standard French (3f, 3m) and six regionally homogeneous speakers of High German from the Saarland (3f, 3m) produced 6 repetitions of the sentences and their *dada* versions from a PowerPoint presentation in response to the recorded questions.

Recordings were made in a sound-treated studio on a Tascam DA-P 1 DAT recorder using an AKG C420^{III}PP headset and transferred digitally via the optical channel to a PC using the Kay Elemetrics MultiSpeech speech signal processing program. Segmentation and labeling with SAMPA was done using the Kiel

XASSP speech signal analysis package, while praat scripts were used for all further signal processing.

1. Durations were calculated for all *feet* in the sentences (German feet are left-headed, French feet right-headed), for the *CW*s, their stressed *syllables* and *vowels*. All durational measurements were normalized relative to the mean duration of the corresponding unit in the sentence.

2. Comparisons focused on the parameter changes in *identical* words across accent levels. Therefore, since segmental structure was identical, F_0 was calculated as the mean fundamental frequency (Hz) across the syllable nucleus of the lexically stressed syllable of CWs, and as F_0 change (contour) by comparison with the unstressed syllables preceding and following them. These values were also normalized by expressing them as percentages of the mean overall F_0 of the sentence.

3. Signal strength was captured in two ways: (i) as the normalized mean intensity (dB) of the syllabic nuclei in the lexically stressed syllable of the CW, expressed as sentence intensity (dB) – syllable intensity (dB), and (ii) as spectral balance, calculated as the difference in intensity between a 70 Hz to 1 kHz and a 1.2 to 5 kHz frequency band of the syllabic nuclei in the lexically stressed syllable of the CW.

4. Spectral definition was captured with the frequency values for formants 1-3 in the middle of the syllabic nucleus in the lexically stressed syllable of CWs.

3. Results

3.1. Language differences

As a first step towards specifying the differences reported in [1] between the accenting and de-accenting patterns in French and German, multivariate ANOVAs were carried out for CW1 and CW2 separately, with language (Fr, G), accent-level (accented, de-accented) and number of syllables in the CW (1,2) as independent variables. We present results for the dada material since this allows direct comparison across the two languages without distortion from different syllable structures.

Table 1: *Main effects for language (French vs. German) (* p<0.05; ** p<0.01; *** p<0.001)*

Parameter	Level of significance	
	CW1	CW2
vowel duration	*	**
syllable duration	***	***
word duration	***	***
foot duration	-	***
F_0 mean	***	-
F_0 difference	***	-
intensity	***	*
spectral balance	***	***
F1	***	***
F2	-	**
F3	*	-

Table 1 shows the main effects for language (over different degrees of accentuation and 1- and 2-syllable CWs) for CW1 and CW2. These main effects indicate that the two languages behave differently with regard to their normalized duration and F_0 values, their energy (dB) difference values (i.e. difference from the sentence mean, or difference between frequency bands), and in their formant values. While this is important to register, the link between these differences and the accenting and de-accenting process in speech production is only indirect – namely, by virtue of the fact that the data reflect the mean and variance of the parameters in the CWs, which have been produced in a context defined as "de-accented" and "accented". More important for the issue addressed in this study are the language × accent-level interactions. These reflect the parameters that are exploited differently in the process of accenting and de-accenting. Table 2 shows the interactions for CW1 and CW2.

Table 2: *Interactions for language (French vs. German) × degree of accent (* p<0.05; ** p<0.01; *** p<0.001)*

Parameter	Level of significance	
	CW1	CW2
vowel duration	**	***
syllable duration	*	***
word duration	*	***
foot duration	-	-
F_0 mean	***	***
F_0 difference	***	***
intensity	-	-
spectral balance	-	-
F1	-	*
F2	**	*
F3	-	-

These results strongly support the differences in the degree of change reported in [1]. For CW1, and even more clearly for CW2, Fr. and G. differ significantly in the degree to which they employ duration for accent differentiation. The differences in the use of F_0 (mean F_0 and F_0 change) which was reported for the *dada* material is also confirmed statistically here. There is no difference in the use of energy measures (mean vowel intensity (dB) and spectral balance) despite the main language effects.

An additional finding, not apparent in the data in [1], is that the two languages also differ in the manner in which the spectral definition of the vowel (the change in the quality as reflected in the formant values) changes between the de-accented and the accented condition. For CW2, i.e., in the differentiation of post-nuclear de-accented and nuclear accented, the degree of opening (F1) is reduced to a greater degree in Fr. /a/ than in G. /a:/. In both CW1 and CW2, the Fr. /a/ is produced with a more retracted quality (lower F2) when accented; G. shows no significant change as a result of (de-) accentuation.

These differences between the languages in the degree of change do not mean that these parameters are not still used by both languages systematically for (de-)accentuation purposes. In fact, when considered separately, *both* languages show significant differences for *all* durational, F_0-linked and energy-linked measures in their production of the de-accented vs. accented CWs. Additionally, F1 and F2 and F3 differ significantly for CW1 in Fr., as do F1 and F2 for CW2. In G., vowel quality parameters vary less with accentuation; only F1 and F2 for CW1 and F1 for CW2 show a significant difference.

Figures 1-4 illustrate the degree to which syllable and word durations change as a function of accentuation. In CW1 (Figs. 1-2), the increase in syllable duration with accentuation is greater for Fr. than G. in all six sentences, and word duration in all but sentence 5. In CW2 (Figs. 3-4), the greater durational change with accentuation is very clear for all six sentences at both syllable and word level.

However, the durational data in the figures also clearly reflect another difference between the temporal structuring of Fr. and G., one which is attributed to the traditional rhythm-typological difference of "syllable-timing" vs. "stress-timing".

The CW1s of both languages comprise 3 monosyllabic and 3 disyllabic words (see section 2.1 above). Figure 1 reveals that the syllable durations of monosyllabic words in G. (sentences 1-3) are significantly longer than the stressed syllable of the disyllables, whereas the slight difference between the monosyllables (sentences 2, 5 and 6) and the stressed syllable of the disyllables (1, 3, 4) are not significant.

The converse is true at word level (Figure 2): In G. there is no systematic difference between the normalized word-durations of the monosyllables and the disyllabic words, whereas the Fr. disyllables are considerably, and significantly, longer than the monosyllables.

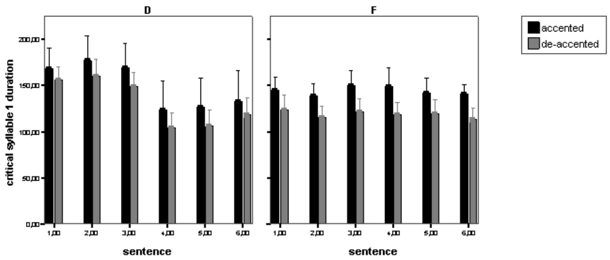

Figure 1: *CW1 syllable duration for German (left tableau) and French (right tableau), sentences 1-6.*

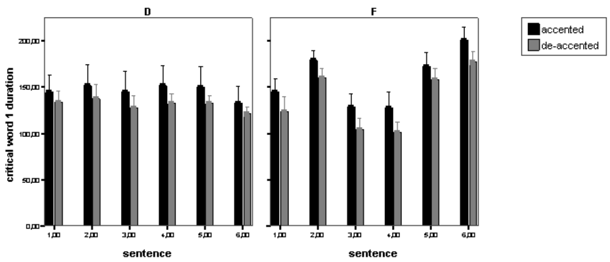

Figure 2: *CW1 word duration for German (left tableau) and French (right tableau), sentences 1-6.*

These observations confirm the tendency for a "syllable-timed" language (Fr. in this case) to maintain a more constant syllable duration independent of the stress status of the syllable, or – as we see here – of the number of syllables in the word. Conversely the tendency for "stress-timed" languages to reduce stressed-syllable duration in polysyllabic words and to reduce the duration of unstressed syllables is borne out by the G. data.

While the "syllable timing" character is just as clear for Fr. CW2 (compare word durations for disyllables in sentence 3, 5 vs. monosyllables in 1, 2, 4, 6), the clear "stress-timing" tendency found for G. CW1 is not so apparent in G. CW2 (compare word durations for disyllables in sentence 1, 2, 6 vs. monosyllables in 3, 4, 5). This, we conjecture, is due to the position of the G. CW2s in the final foot of the sentence in all except sentence 6, whereas the Fr. CW2s are all in the penultimate foot. If the domain of final lengthening is the foot, then the second (unstressed) syllable of disyllabic words will be subject to lengthening, thus distorting the monosyllable-disyllable equality found in CW1.

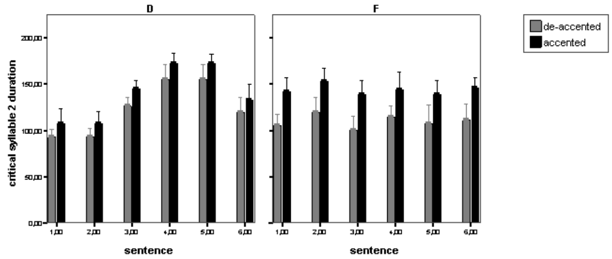

Figure 3: *Critical syllable 2 duration for German (left tableau) and French (right tableau), sentences 1-6.*

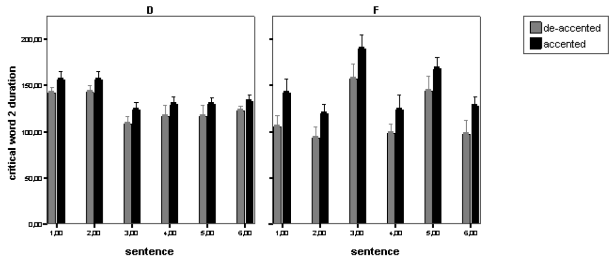

Figure 4: *Critical word 2 duration for German (left tableau) and French (right tableau), sentences 1-6.*

Interestingly, there is no difference between the languages in the modification of foot duration for accentuation (cf. Table 2), though the foot is generally seen as the key unit in the definition of stress-timing. However, the structure of the speech material examined here is not suited to exploring the word-foot relationship thoroughly in terms of their role in accentuation.

In summary, the individual language ANOVAs show that Fr. and G. both use the same parameters for accentuation differentiation. At the same time there are many main effects of language as well as significant interactions between language × accent-level to indicate differences in the degree of exploitation. In the following section, we therefore examine further the *differential* contribution of those parameters to the prediction of accent class (de-accented or accented) by means of discriminant analyses.

3.2. Accent prediction

The method employed was a *stepwise* discriminant analysis carried out separately for the two languages, and for text and reiterant ”*dada*” replies to the questions. As in the previous section, only the results for reiterant replies are discussed here, since they allow a direct comparison of the two languages unobscured by the variability in the segmental structuring of the CWs.

The criterion variable ‘accentuation’ discriminates between accented and de-accentuated. All parameters listed in Section 2.2 were used as predictor variables. The contribution of the individual parameters to the discrimination between these two categories is derived from the standardized discriminant coefficients (sdc) in the final discriminant model. Correct classification varied between 98.1-100%.

The results are summarized in Table 3. In both Fr. and G., F_0 and intensity are the main predictors for discrimination of accented versus de-accented CWs, independent of their position in the utterance (early or late). But note that the order of the two parameters, reflecting their relative weighting, is reversed in the two languages, with F_0 being more important for Fr., while intensity is more important for G. Another striking difference is the role of spectral balance in G., even though its weight is much less than that of F_0 and intensity. It does not enter into the discriminant function for Fr. at all.

Table 3: *Standardized discriminant coefficients (sdc's) for French and German reiterant CWs*

CW	Language	Parameter	sdc
CW1	French	mean F_0	0.709
		syllable duration	0.665
		vowel duration	-0.379
		intensity	0.328
CW1	German	intensity	0.683
		mean F_0	0.575
		word duration	0.399
		spectral balance	-0.209
		vowel duration	0.171
		foot duration	0.158
CW2	French	mean F_0	0.962
		intensity	0.576
		F_0 change	-0.419
		vowel duration	0.279
		word duration	0.164
CW2	German	Intensity	0.932
		vowel duration	0.671
		mean F_0	0.515
		syllable duration	-0.430
		spectral balance	-0.345

Spectral definition parameters are never selected as predictors for accentuation discrimination, despite our previous observation (Section 3.1). The results for duration, which has been found to be of great importance in so many studies on stress and accent and which was shown here to differ in Fr. and G. (cf. the MANOVA results in section 3.1), are difficult to interpret. Duration clearly *does* play a role, but for both languages with comparable sdc values. However, there is no particular (set of) duration parameter(s) which is picked out as the best predictor for one or the other language. This is presumably due to the high level at which all the duration parameters correlate with one another.

Nonetheless, the (partial) overlap between the languages in the predictors selected for accentuation differentiation is in agreement with our MANOVA results, and underlines that we are examining the "universal" phonetic means of signaling prominence. The variable weights of different predictors show that differences between languages are "more-or-less", not "all-or-none". Therefore the considerable degree of co-variation between parameters is to be expected.

The strength of the contribution of the modestly varied intensity parameter and the (relatively uncorrelated) F_0 parameter to accent prediction and the confusing picture of the contributions made by the large durational changes relativizes the usefulness of discriminant analysis methods. Intensity changes are small but consistent across all structural differences, whereas duration is sensitive to syllable and word structure. This increases the within-factor variation, reducing the apparent contribution of the accentuation-linked changes.

4. Discussion

The differences between the languages in relation to their exploitation of the acoustic dimensions duration, F_0, signal strength and spectral definition can be only partially explained with reference to differences in the phonological structure of the two languages.

The differences in the degree to which accented vowels and syllables are lengthened in Fr. compared to G. (compare [1] and the MANOVA results reported above) may be seen in the context of vowel-length oppositions in G. constraining the exploitation of duration.

There may also be a link between the inclusion of spectral balance as an accent predictor for G. in contrast to Fr. (cf. Table 3). As a so-called 'stress-timed' language, G. can be expected to differentiate the production effort for non-accented vs. accented syllables to a greater extent than Fr. However, this finding needs to be treated with caution, since although there is a main language effect for spectral balance (cf. Table 1), there is no language × accent-level interaction (Table 2) to suggest its systematic differential exploitation for accent production. Also, the expected accompanying change – a spectral shift in de-accentuated vowels, which is significantly different for the two languages (cf. Table 2) – is in fact stronger for Fr. than G. (F1, CW2, Δacc.–deacc.: Fr. 141 Hz, G. 74 Hz.; F2, CW1, Δacc.–deacc.: Fr. –88 Hz, G. 24 Hz; CW2, Δacc.–deacc.: Fr. –94 Hz, G. 6 Hz). The reversed predictor weighting in Fr. and G. for the two main predictors of accentuation in the discriminant analysis, mean F_0 and intensity, has no apparent grounding in the phonologies of the two languages. Both languages exploit F_0, but they also clearly use it differently (cf. Table 2). A detailed intonation analysis is clearly necessary, and the parameterization of F_0 difference may need refinement to capture not only the degree but also the type of change. The result then may well be a number of differences in prominence-giving production behaviour linked to the intonational phonology of the languages.

6. References

[1] Barry, W.J.; Andreeva, B.; Steiner, I. The Phonetic Exponency of Phrasal Accentuation in French and German. *Proc. Interspeech,* Antwerp, 2007.

[2] Fry, D.B. Experiments in the perception of stress. *Lang. & Speech* 1, 126-152, 1958.

[3] Fry, D.B. The dependence of stress judgments on vowel formant structure. *Proc. 6th ICPhS,* Münster, 306-311, 1965.

[4] Kochanski, G., Grabe, E., Coleman, J. Loudness predicts prominence: fundamental frequency lends little. *J. Acoust. Soc. Amer*. 118, 1038-1054, 2005.

[5] Sluijter, A., van Heuven, V. Spectral balance as an acoustic correlate of linguistic stress. *J. Acoust. Soc. Amer*. 100, 2471-2485, 1996.

[6] van Kuijk, D., Boves, L. Acoustic characteristics of lexical stress in continuous speech. *Sp. Comm*. 27, 95-111, 1999.