

TEMPORAL-ALIGNMENT CATEGORIES OF ACCENT-LENDING RISES AND FALLS

David House*, Dik Hermes** and Frédéric Beaugendre**+

*Department of Linguistics and Phonetics, Lund University, Helgonabacken 12, S-223 62 Lund, Sweden

E-mail: david.house@ling.lund.se

** IPO, Center for Research on User-System Interaction, P.O. Box 513, NL 5600 MB Eindhoven, The Netherlands

E-mail: hermes@ipo.tue.nl, frederic.beaugendre@lhs.be

+Currently with Lernout & Hauspie Speech Products

ABSTRACT

This paper presents a theoretical framework for a model of perceived accentuation categories. This framework is based on the combined results of a series of experiments on accentuation boundaries in Dutch, French and Swedish and on theoretical work on tonal perception in speech. We propose a model in which several different language-dependent categories of accentuation are represented as the falling or rising pitch movement is advanced through the syllable. The perceived category depends upon whether an onset of the pitch movement or a pitch jump is perceived, whether or not the particular category is represented in the language in question, and whether or not the movement also serves as a cue for phrasing. Dutch and Swedish display similarities in accentuation categories while French differs from these two languages. These differences are explained by the conflict between cues for accentuation and phrasing and the differing intonational structure of French. The proposed perceptual categories have general implications for the understanding and description of accentuation.

1. INTRODUCTION

Considerable attention has been directed recently at the temporal-alignment characteristics of accent-lending pitch movements in terms of the onset of the stressed vowel. For example, in the description of Dutch intonation in [13], two kinds of accent-lending rises are distinguished: an early rise starting before the vowel onset and a late rise starting after the vowel onset. These categories correspond with $L+H^*$ and L^*+H in autosegmental terminology [3]. In Swedish, the rise is coupled to focal accent and is phonologically separate from the two different word-accent falls which are described as $H+L^*$ (acute accent) and H^*+L (grave accent) (Bruce [4]). For Dutch, 't Hart et al. [13] present only one phonetic category of full-sized, accent-lending falls, while Rietveld and Gussenhoven [11] distinguish two phonological categories H^*L and $!H^*L$.

In recent work on the timing of pitch movements and perception of accentuation in Dutch, Hermes [6,7] established that the main cue which induces the percept of accentuation is the onset of a pitch movement. By shifting the timing of pitch movements (both rises and falls) in /mamamama/ and /aaaaa/ utterances from left to right, he determined the point for onsets of pitch movements where the percept of accentuation shifts from one syllable to the next and referred to it as the *accentuation boundary*. For Dutch, the accentuation boundaries were located near the end of the /a/ vowel near the following syllable onset.

These results led to the question of language specific accentuation boundaries. Are the boundaries universal, or do they differ in relationship to accentuation structure? To address this question, a set of similar experiments was carried out using French and Swedish subjects.

2. METHOD

A short summary of the methodology is presented here. The different experiments are presented in more detail in [2,6,7,8]. The experiments were carried out with reiterant five-syllable utterances of the type /mamamama/ and /aaaaa/. The middle three syllables were exact replicates of each other as to amplitude and spectral content. The original pitch contour was replaced by a rising or a falling pitch movement superimposed on a declination line using pitch-synchronous-overlap-add (PSOLA) techniques.

The timing of the pitch movements, both rises and falls, was systematically varied in such a way that the movements accented the third or the fourth syllables (second or third for the Dutch stimuli). These two syllables are referred to here as *test syllable one* and *test syllable two*. This was done for pitch movements of three different durations (80, 120 and 160 ms) and for three different durations of the /m/ in the /mamamama/ stimuli and the silence between the /a/s in the /aaaaa/ stimuli (i.e. short, normal and long). Subjects were asked to indicate which syllable they perceived as accented. The accentuation boundary was then defined as the

moment before which more than half of the subjects indicated the previous syllable as accented and after which more than half of the subjects indicated the next syllable as accented.

3. RESULTS

The results for both French and Swedish confirm the finding for Dutch that the cue which induces the percept of accentuation is located at the onset of the pitch movement. Furthermore, Figure 1 shows response distributions of the onset times for the 80 ms rises and falls accenting /ma/ test syllable one for Dutch, Swedish and French subjects. For the rises, the accentuation boundary is located at the end of the vowel for Dutch and Swedish listeners, while for French listeners, the boundary is located earlier, near the middle of the vowel. For the falls, the boundary is also located at the end of the vowel for Dutch and Swedish listeners. For the French listeners, however, the boundary for the falls was not well defined. Similar results are presented for the /a/ test syllables in Figure 2.

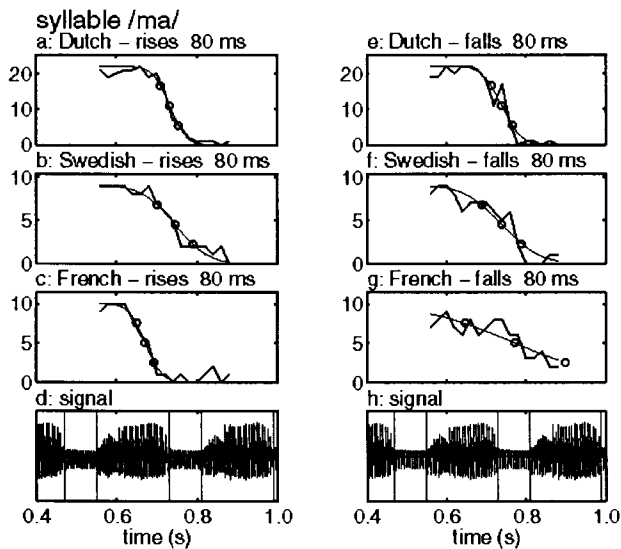


Figure 1. Response distributions (thick line) of the onset times for the 80 ms rises and falls accenting /ma/ test syllable one for Dutch, Swedish and French subjects. The thin line indicates an error function fitted through these distributions. The circles indicate the quartiles of these error functions.

In certain cases there is also an area of ambiguity at the boundary even for the Dutch and Swedish listeners which shows up as a plateau in the response distributions. This plateau is clearly present in the results for Dutch subjects for the

falls of different durations in the /a/ syllable in the long-silence condition as presented in Figure 3.

Perhaps, a plateau can also be observed for the rises, but it is lower than for the falls. This may indicate that the stimuli with the rises are less ambiguous than those with the falls.

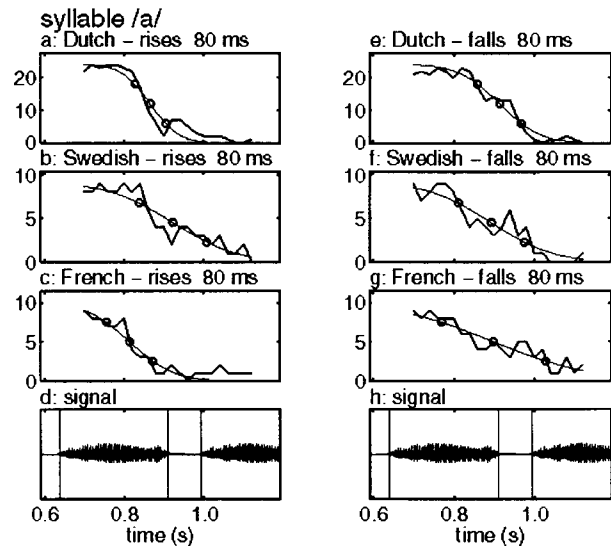


Figure 2. Response distributions of the onset times for the 80 ms rises and falls accenting /a/ test syllable one for Dutch, Swedish and French subjects. The circles indicate the quartiles of the error function fitted through these distributions.

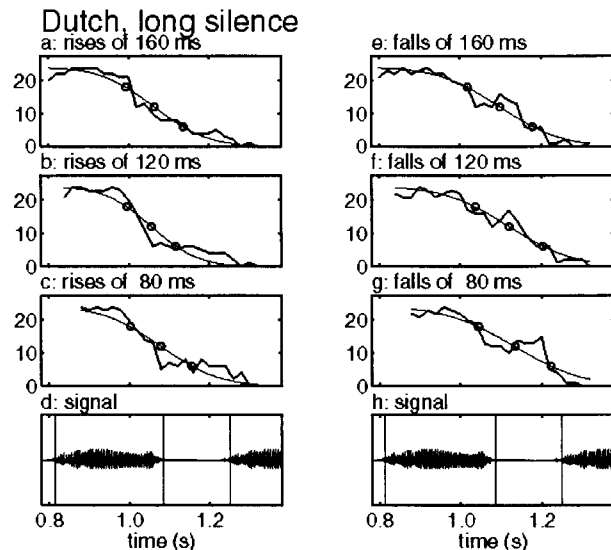


Figure 3. Response distributions of the onset times for the 80 ms, 120 ms, and 160 ms rises and falls accenting /a/ test syllable one (long silence condition) for Dutch subjects. The circles indicate the quartiles of the error function fitted through these distributions.

4. DISCUSSION

First of all, the fact that the responses of the French subjects differ from those of the Dutch and Swedish subjects is not particularly surprising as accentuation in French differs considerably from accentuation in Dutch and Swedish. French accent is often defined in the literature as a fixed accent: its occurrence is not imposed by morphology (there are no variable distributions used to differentiate lexical categories); it is often said that French accent is always placed on the last syllable of a lexical word and for that reason it is sometimes called word accent. Actually, different works have underlined the inadequacy of such a definition: French accentual structure is a more complex structure composed of several accents with different functions and distributions as described in Rossi [12]:

Primary accent: In French, the primary accent (or internal accent) is used to group words which are strongly linked at the syntactic level into one stress unit. This syntactic function explains the occurrence of the primary accent at the end of a constituent. The perceptual cues for the primary accent are duration and pitch.

Secondary accent: other types of accents are often grouped under the name of secondary accent [10]. However, in Rossi [12], different types of “secondary” accents are distinguished, along with a specific function and domain: - accent of focalization (or emphatic accent): this accent is used to focus an element within a phrase [14]. One particular manifestation of the accent of focalization is a demarcative accent which occurs systematically at the beginning of the accentual group; its perceptual cues are pitch and loudness; - melodic ictus: this accent usually occurs on the first syllable of words. Its function is merely rhythmical. Its perceptual basis is pitch. A number of studies have shown where those different accents are distributed within the sentence and the syntactic or prosodic group [1,5,10,12,16].

However, only a small number of these studies investigated the timing of the accent in relationship with vowel onset [1,15]. Those studies claim that falls do not have a clear function in the accentual structure of French, whereas rises define the melodic variation for a primary accent as a rise on the whole syllable and the melodic variation for a secondary accent as a rise on the consonant preceding the accented vowel. The fact that no other timing is considered for the secondary accent for French can be interpreted as indicating that French listeners do not normally make the distinction between different timings of movements. Our results clearly confirm that French listeners do not

classify as accented a syllable with a falling pitch movement.

Taken together, the results from the three languages suggest a model of perceived accentuation in which several different categories of accentuation are represented as the falling or rising movement is advanced through the two test syllables. The perceived category depends upon whether an onset of the pitch movement or a pitch jump is perceived [6,7,9] whether or not the particular category is represented in the language in question, and whether or not the movement also serves as a cue for phrasing. A jump category will be perceived when the pitch movement occurs through spectral change at the syllable onset or simultaneously with vowel onset. The perceived onset category will be perceived when the pitch movement occurs through relative spectral stability in the vowel [9].

For rises, four categories can be described, two linked with test syllable one and two linked with test syllable two. This will be formalized in autosegmental terminology [3] as follows:

1. V1 perceived jump category $L+H^*$ = low plus high (V0 to V1)
2. V1 perceived onset category L^*+H = rise (on V1)
3. V2 perceived jump category $L+H^*$ = low plus high (V1 to V2)
4. V2 perceived onset category L^*+H = rise (on V2)

where V1 is the first possible vowel perceived as accented (the vowel in test syllable one) and V2 is the second possible vowel perceived as accented (the vowel in test syllable two).

The language differences for the rises can be explained if the French listeners are mainly using the two categories 1 and 3 (the jump categories). From the perceptual point of view this would mean that in French the onset cue has much less weight than it has in Dutch and in Swedish. This would put the boundary earlier in the vowel in the middle of what is category 2 for the Swedish and Dutch listeners who in turn have their category boundary between 2 and 3. For Dutch, categories 1 and 3 correspond with the early accent-lending rise, while categories 2 and 4 correspond with the late accent-lending rise as distinguished in the IPO description of Dutch intonation [13].

The situation for the falls is more complex in that in some positions the falls can induce a strong percept of phrasing which then seems to override the percept of accentuation. This situation especially occurs when the interval between the

offset of the vowel in test syllable one and the following vowel onset of test syllable two is long. Here six categories are proposed, four inducing accentuation and two inducing phrasing. Perceived phrasing categories produce ambiguous responses which show up as the plateaus in the response distributions.

The categories are formalized as follows:

1. V1 perceived jump H% L (V0 to V1) no perceived accent (boundary cue)
2. V1 perceived jump H+L* = high plus low (V0 to V1)
3. V1 perceived onset H*+L = fall (on V1)
4. V2 perceived jump H% L (V1 to V2) no perceived accent (boundary cue)
5. V2 perceived jump H+L* = high plus low (V1 to V2)
6. V2 perceived onset H*+L = fall (on V2)

If the French listeners have only categories 1 and 4, this would explain their difficulty in hearing the fall as an accent in any position. The Dutch/Swedish listeners hear 2 and 3 and 5 and 6 as accentuation while 1 and 4 would explain the areas of ambiguity for accentuation. Finally the H+L* and H*+L distinction corresponds to the two Swedish word accents for Swedish [4] and two categories of accent-lending falls in Dutch. Categories 1 and 4 would correspond with the non-accent-lending fall in the IPO description of intonation [13], while categories 2, 3, 5 and 6 correspond with the accent-lending fall, which are not further distinguished in the IPO approach.

The perceptual cues for these distinctions are provided by the location of the onset of the pitch movement, the jump in pitch from one vowel to the next including its direction, and the interval between the end of the vowel and the following vowel onset. The differences between the languages as well as the differences between rises and falls can be explained by attributing different perceptual weights to these cues.

5. REFERENCES

- [1] Beaugendre F. *Une étude perceptive de l'intonation du français*. Thèse de doctorat de l'université Paris XI, Orsay, 1994.
- [2] Beaugendre, F., House, D., and Hermes, D.J. Accentuation boundaries in Dutch, French and Swedish. Accepted for *Proc. of the ESCA Tutorial and Research Workshop on Intonation*, 18-20 September, 1997, Athens, Greece.
- [3] Beckman, M.E. and Pierrehumbert, J.B. Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255-309, 1986.
- [4] Bruce, G. *Swedish word accents in sentence perspective*. Gleerup, Lund, Sweden, 1977.
- [5] Fonagy I. L'accent français: accent probabilitaire; L'accent en français contemporain. *Studia Phonetica* 15, 123-233, 1979.
- [6] Hermes, D.J. Timing of pitch movements and accentuation of syllables. *Proc. ICSLP 96, Fourth International Conference of Spoken Language Processing*, pp. 1197-1200, Philadelphia, PA, USA, 1996.
- [7] Hermes, D.J. Timing of pitch movements and accentuation of syllables in Dutch. Submitted for publication, 1997.
- [8] Hermes, D.J., Beaugendre, F., and House, D. Temporal alignment of accentuation boundaries in Dutch. Accepted for *Proc. of the ESCA Tutorial and Research Workshop on Intonation*, 18-20 September, 1997, Athens, Greece.
- [9] House, D. *Tonal Perception in Speech*. Lund University Press, Lund, Sweden, 1990.
- [10] Padeloup V. *Modèles de règles rythmiques du français appliqué à la synthèse de la parole*. Thèse de 3ème cycle, Institut de Phonétique d'Aix-en-Provence, 1990.
- [11] Rietveld T. and Gussenhoven C. The influence of syllable composition on the alignment of pitch targets. *Journal of Phonetics* 23, 375-385, 1995.
- [12] Rossi M. A model for predicting the prosody of spontaneous speech (PPSS model). *Speech Communication* 13, 87-107, 1993.
- [13] 't Hart, J., Collier, R., and Cohen, A. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge University Press, Cambridge, UK, 1990.
- [14] Touati, P. *Structures prosodiques du suédois et du français*. Lund University Press, Lund, Sweden, 1987.
- [15] Vaissière J. La structuration acoustique de la phrase française. *Annali della scuola normale superiore di Pisa*, pp. 530-560, 1980.
- [16] Verluyten S.P. Phonetic reality of linguistic structures: the case of secondary stress in French. *Proceedings of the 10th International congress in phonetic sciences*, Utrecht, pp. 522-526, 1984.