



HAL
open science

Segmental “Anchorage” and the French late rise

Pauline Welby, H el ene Loevenbruck

► **To cite this version:**

Pauline Welby, H el ene Loevenbruck. Segmental “Anchorage” and the French late rise. Interspeech 2005: the 9th Annual Conference on Speech Communication and Technology, Sep 2005, Lisboa, Portugal. pp.2369-2372. hal-00371967

HAL Id: hal-00371967

<https://hal.science/hal-00371967>

Submitted on 30 Mar 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d’enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

Segmental “anchorage” and the French late rise

Pauline Welby, H el ene L aevenbruck

Institut de la Communication Parl ee, CNRS UMR 5009
Institut National Polytechnique de Grenoble, Universit e Stendhal, Grenoble, France
{welby, loeven}@icp.inpg.fr

Abstract

We examined the tonal alignment and scaling patterns of the start and end points of the French late rise, using a rate manipulation paradigm. Our findings call into question aspects of the segmental anchoring hypothesis: the low starting point of the late rise was not stably anchored to a segmental landmark, and for some speakers, F_0 excursion size varied across rates. The position of the peak of the late rise was found to vary across syllable structures. To account for the observed patterns, we propose the notion of an “anchorage,” that is, a region within which an intonational turning point can anchor. For the peak of the French late rise, this anchorage stretches from just before the end of the vowel of the last full syllable of the accentual phrase to the end of the phrase.

1. Introduction

1.1. Segmental anchoring

The “segmental anchoring” hypothesis makes strong predictions about the nature of fundamental frequency (F_0) rises and falls cross-linguistically. It claims that the temporal alignment of both the start and the end of an F_0 movement will be defined with respect to landmarks in the segmental string ([1], [2]), and studies report a remarkable precision of the start and end of F_0 rises and falls with respect to these “anchor” points, sometimes on the order of a few milliseconds (see, for example, [3]).

Proponents of segmental anchoring have also assumed that F_0 excursion size will remain fairly stable. This tonal scaling assumption is crucial to a rate manipulation paradigm used in segmental anchoring studies (e.g., [4]). Ladd et al., 1999 ([1]), who introduced the term “segmental anchoring,” reasoned that given stable temporal alignment and stable F_0 excursion size, F_0 rises in fast speech should have shorter rise times and thus steeper slopes. They tested this hypothesis against a competing hypothesis in which speakers seek to maintain a constant slope.

1.2. Some basics of French intonational structure

All accounts of French intonation agree that the utterance is divided into smaller units, called *Inton eme intonatif mineur* ([5]), Accentual Phrase (AP) ([6]), etc. by various researchers. This AP is typically characterized by an F_0 rise on the last syllable of a phrase that is not utterance-final and an optional early rise near the beginning of the phrase. The late (final) rise is a marker of the “primary accent,” and the early (initial) rise is a marker of the “secondary accent.” In an Autosegmental-Metrical framework (AM, e.g., [7]), an AP with this two-rise pattern would be transcribed with a series of L(ow) and H(igh) tones, LHLH. An example, drawn from the current study, is shown in Fig. 1.

1.3. Goals of the study

The first main goal of the study was to examine the claims of segmental anchoring with respect to the start (L2) and end (H2)

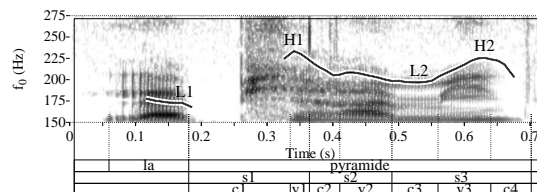


Figure 1: Accentual phrase *la pyramide* produced with an early and a late rise (LHLH). Target syl.: /mid/, (CVCobs).

of the French late rise.

The second goal was to determine the precise landmark to which H2 is aligned, a question left open in our earlier studies, in which we concluded (rather vaguely) that this peak was achieved “close to the end” of the last syllable of the AP ([8]).

1.3.1. Syllable structure and H2 alignment

We examined whether H2 alignment differs according to whether the associated syllable (the last full syllable) is open (CV, e.g., *salami* /sa.la.mi/), closed by a sonorant (CVCson, e.g., *vitamine* /vi.ta.min/) or closed by an obstruent (CVCobs, e.g., *pyramide* /pi.ra.mid/). This allowed us to examine three potential “segmental anchors”: (1) end of the last syllable vowel, (2) end of the sonorant rhyme, (3) end of the syllable. Each of these makes different predictions about H2 alignment. In the first case, H2 should be systematically realized at a fixed position with respect to the end of the vowel, regardless of rhyme structure. In the second case, we expect H2 to be realized at a fixed position with respect to the vowel in CV and CVobs items, but with respect to the coda consonant for CVCson ones. In the last case, we expect the same pattern of temporal alignment, but with possible undershoot of H2 in CVCobs syllables.

1.3.2. L2 alignment

Our earlier work showed considerable variability in L2 alignment: L2 was sometimes realized in the final syllable of the AP, sometimes in the penultimate syllable ([8], [9], see also [10]). It was thus not possible to suggest plausible specific segmental anchors *a priori*. We instead examined whether L2 was associated to one of the last two syllables of the AP or whether L2 behaved like a “leading tone” of H2 in the usual Autosegmental-Metrical meaning of this term. If the former is true, we expect to find that L2 is consistently aligned with respect to a syllable landmark, and to be able to identify potential segmental anchors. If the latter is true, we expect that L2 will be realized at an “fairly invariant” interval before H2 ([7]) and that rise times will vary very little within a speaker. Even if no invariant interval between L2 and H2 is found, if both are consistently realized near the end of the AP, this will be evidence for an alternative hypothesis in which L2 and H2 form a loose unit.

1.3.3. Hypotheses concerning the L2/H2 relationship

We also examined a number of hypotheses concerning the relationship between L2 and H2, notably whether a constant rise time, F_0 excursion size, or slope is found within a given rate and under changes in rate.

2. Methods

2.1. Materials

The corpus contained six short paragraphs with triplets of target words.¹ An example is given in (1), with targets underlined.

- (1) La cuisine de Marie-Noëlle regorge de victuailles. Un salami qui a été rapporté de Parme suinte d'une graisse très prometteuse. Sur la desserte, un immense compotier déborde de fruits de toutes sortes. Les vitamines qu'ils recèlent naturellement sont un prétexte pour les gourmands. Une pièce-montée en chou et nougatine trône sur le buffet. La pyramide qui dégouline de caramel fera sûrement des amateurs.

'Marie-Noëlle's kitchen is overflowing with provisions. A salami that was brought back from Parma is dripping with very promising grease. On the side table, a giant bowl is brimming with all kinds of fruit. The vitamins that they naturally contain are just an excuse for people who like to eat. A layer cake with puff pastries and nougatine is prominently displayed on the cabinet. The pyramid, dripping with caramel, will surely be appreciated.'

Within each triplet, the target words differed in the rhyme structure of the last syllable: it was either CV (e.g., *salami*), CVCson (e.g., *vitamine*), or CVCobs (e.g., *pyramide*). All targets were three-syllable nouns or substantive adjectives with the syllabic structure $C_1V_1.C_2V_2.C_3V_3(C_4)$, where C_3 was a sonorant, and to the extent possible, C_1 and C_2 were at least voiced. Within a triplet, the sequence $V_2.C_3V_3$ in the target region (e.g., /a.mi/ in the triplet in (1)) was identical.

Each target word appeared sentence-initially, preceded by a form of the definite article (*le, la, les*). The resulting noun phrase was followed by a relative clause beginning with a relative pronoun (*qui, qu'il, qu'elle, qu'ils*), which was followed by the verb phrase of the main clause.

The targets contained a variety of vowels (/a, ε, i, o/), sonorant coda consonants (/m, n, l/) and voiced and voiceless obstruent coda consonants (/d, z, t, s/). Voiced coda obstruents, however, were always at least partially devoiced.

2.2. Participants

The participants were six native speakers of standard Hexagonal French,² all women from the Paris region. They were all students or researchers at universities in Grenoble and had an average age of 26.8. Five of the speakers were naive to the hypotheses; Speaker 6 is author HL. Speakers were paid €10 for their participation, which lasted about a half hour.

2.3. Procedures

Each participant read the six paragraphs, first at a self-selected normal rate, then at a fast rate. Speakers were recorded at 44.1 kHz onto digital audio tape (DAT), using a Shure SM10A head-worn microphone in a sound attenuated chamber at the Institut de la Communication Parlée. The data were transferred to computer, then downsampled to 22.05 kHz. The soundfiles were segmented and each utterance saved as a separate file. F_0 curves and spectrograms were created using Praat speech analysis software ([11]).

¹A second corpus, a list of sentences, was also recorded. Due to space constraints, we discuss only the paragraph corpus here. The same general patterns of alignment held across both corpora.

²The French of France, l'Hexagone.

For the target word in each utterance, word, syllable, and segment boundaries were tagged, as were the syllables immediately preceding and following the target and the beginning and the end of the utterance, using waveforms and spectrograms to guide the segmentation. Praat scripts were written to semi-automate this process.

The intonational features tagged included H1 (peak of early rise), L2 (start of late rise), and H2 (peak of late rise) (see Fig. 1). The position of peaks was defined in the labelling script as the time of the local f_0 maximum, automatically detected, then hand-corrected if necessary. Since L2 did not always correspond to an F_0 minimum, a line-fitting procedure was used to automatically calculate the position of L2. We do not detail the procedure here, but refer the reader to [8] and [9] for a detailed explanation.

3. Results

3.1. Rate manipulation

Average speaking rate in syllables/second was calculated for each speaker in each speaking rate. Rate was calculated across entire paragraphs, across sentences, and across target words. All speakers successfully raised their speaking rate from the normal to the fast rate for all three sets of comparisons.

3.2. Rise time, F_0 excursion, slope analyses

For these and subsequent analyses, we examined only items with the LHLH pattern. Of the 216 items (3 target words \times 6 paragraphs \times 6 speakers \times 2 rates), 80 were produced with other patterns, and were excluded. Three items produced with a clear final schwa in the target AP were also excluded. In addition, we do not consider Speaker 6's data for this corpus, since she produced only ten LHLH items. This left a total of 123 items for the analyses.

The statistical analyses were somewhat complicated by the unexpectedly low rate of LHLH patterns in target APs. Comparison of the normal vs. fast rate productions was not possible for all pairs.

We performed paired t-tests between the normal and the fast rate means for F_0 excursion ($H2F_0 - L2F_0$), rise time ($timeH2 - timeL2$), and slope (F_0 excursion/rise time). Means are given in Table 1.

Rise time Rise times were numerically longer in the normal rate than in the fast rate for all speakers, but this difference was statistically significant for only two speakers (Speaker 2: $t = 3.28$, $df = 9$, $p < 0.01$; Speaker 5: $t = 3.02$, $df = 15$, $p < 0.01$). Independent t-tests using all 123 items with LHLH in the target region showed the same pattern for these two speakers.

Speakers did not maintain a constant rise time, as shown

S	rise time (S.E.)		F_0 . exc. (S.E.)		slope (S.E.)		N pairs
	norm	fast	norm	fast	norm	fast	
1	138 (13)	126 (12)	119 (19)	95 (15)	.93 (.23)	.75 (.12)	5
2	127 (11)	95** (7)	69 (9)	48* (4)	.57 (.08)	.51 (.04)	10
3	109 (8)	96 (11)	28 (3)	27 (4)	.26 (.03)	.28 (.03)	7
4	116 (9)	100 (7)	35 (3)	38 (3)	.31 (.02)	.39* (.03)	9
5	114 (7)	91* (7)	30 (2)	31 (3)	.28 (.03)	.38* (.04)	16

Table 1: Mean F_0 excursion (Hz), rise time (ms), and slope (Hz/ms) by speaker and rate. * $p < 0.01$, ** $p < 0.05$.

by the standard errors in Table 1. In the 21 items available for Speaker 1, for example, rise times for the late rise varied from a minimum of 104 ms to a maximum of 167 ms.

F_0 excursion Another series of paired t-tests tested the effect of rate on F_0 excursion size. Analyses were performed in hertz, ERBs ([12]), and semitones ([13]). Values are reported in hertz. Only Speaker 2, showed a significant effect of rate on F_0 excursion, with larger F_0 excursions in the normal rate ($t = 2.49, df = 9, p < 0.05$). The same results were found using semitones and ERBs, and in the independent t-test analyses.

Slope We examined whether speakers seek to maintain a constant slope in two ways. We first performed a series of paired t-tests for each speaker to test whether slope differed across rates. For two speakers, there was a significant rate effect (Speaker 4: $t = 3.11, df = 15, p < 0.01$; Speaker 5: $t = 5.27, df = 8, p < 0.01$). The same pattern was found in the ERB and semitone analyses. In the independent t-test analyses, the same pattern was found for Speaker 5. We then performed correlation analyses to examine whether rise time and F_0 excursion size were positively correlated, as predicted by a constant slope hypothesis (as described by [1]). The correlations were significant for two speakers, Speaker 2 ($r = 0.442, p < 0.05$) and Speaker 3 ($r = 0.527, p < 0.05$).

3.3. H2 scaling in CVCobs vs. CV and CVCson syllables

We compared H2 scaling in syllables with obstruent rhymes vs. those with sonorant rhymes, to see whether there was undershoot of H2 in CVCobs syllables. For only one speaker, Speaker 4, was there a significant difference: H2 was significantly lower in CVCobs syllables than in CV and CVCson syllables (213 Hz vs. 222 Hz; $t = 2.64, df = 23, p < 0.05$).

3.4. L2 alignment

We examined L2 alignment, first plotting the results by speaker and rate. L2 was almost always realized in the last syllable of the accentual phrase, although it was also occasionally realized in the penultimate syllable. We examined L2 alignment with respect to two landmarks: the beginning of the last syllable of the AP (begS3toL2) and the beginning of the vowel of that syllable (begV3toL2). The results are shown in Table 2. Paired t-tests were conducted to examine the potential effect of rate. For two speakers, L2 alignment was affected by rate. For Speaker 1, L2 was aligned later in the normal rate than in the fast rate (begS3toL2: $t = 4.11, df = 4, p < 0.05$, begV3toL2: $t = 3.82, df = 4, p < 0.05$). Speaker 5 showed the reverse pattern (begS3toL2: $t = 2.67, df = 15, p < 0.05$, begV3toL2: $t = 2.65, df = 15, p < 0.05$). Similar patterns were found in the independent t-tests analyses.

S	Norm.	Fast	Norm.	Fast	N pairs
	begS3 toL2 (S.E.)	begS3 toL2 (S.E.)	begV3 toL2 (S.E.)	begV3 toL2 (S.E.)	
1	80 (12)	15* (7)	4 (10)	-44* (9)	5
2	55 (7)	53 (7)	-25 (7)	-20 (6)	10
3	63 (14)	37 (18)	1 (9)	-14 (16)	7
4	80 (12)	80 (13)	3 (14)	10 (13)	9
5	54 (6)	68* (7)	-14 (6)	3* (6)	16

Table 2: L2 alignment (ms). * $p < 0.05$.

S	S3 structure	endV3toH2	endSonRhymetoH2	N
		(S.E.)	(S.E.)	
1	CVCobs	8 (7)		8
	CVCson	15 (8)	-71 (7)	6
	CV	-8 (3)		7
2	CVCobs	-9 (4)		8
	CVCson	31 (10)	-37 (4)	8
	CV	-14 (3)		8
3	CVCobs	-9 (3)		6
	CVCson	22 (9)	-45 (8)	6
	CV	-12 (2)		7
4	CVCobs	-11 (3)		10
	CVCson	38 (6)	-50 (3)	8
	CV	-10 (5)		7
5	CVCobs	-13 (2)		11
	CVCson	34 (5)	-54 (7)	12
	CV	-15 (2)		11

Table 3: H2 alignment (ms).

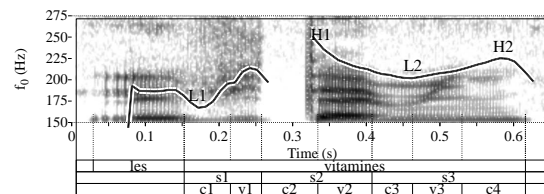


Figure 2: Accentual phrase *les vitamines* produced with an early and a late rise (LHLH). Target syl.: /min/, (CVCson).

3.5. H2 alignment

Plots of H2 alignment by speaker revealed apparent alignment differences based on syllable structure, but not on rate. We therefore combined rates in the alignment analyses. We examined H2 alignment with respect to two landmarks: the end of the last syllable vowel (endV3toH2) and the end of the sonorant rhyme of that syllable (endSonRhymetoH2). As shown in Table 3, the following general patterns of H2 alignment were found. First, H2 was generally aligned later in CVCson syllables than in CV and CVCobs syllables (cf. H2 alignment, Fig. 2 vs. Fig. 1). Second, in CVCobs and CV syllables, H2 was realized within 20 ms of the end of the vowel. Third, H2 in CVCson syllables was generally realized before the end of the coda consonant (as shown by the negative endSonRhymetoH2 means). Finally, standard error tended to be larger for the CVCson structure, indicating greater variability in H2 alignment.

For Speakers 2, 3, 4 and 5, all comparisons between CVCson and the other two structures (and only those comparisons) were significant, for both endV3toH2 and endSonRhymetoH2 comparisons. For Speaker 1, the following comparisons were significant: CVCson vs. CV for endV3toH2, CVCson vs. CVCobs and CVCson vs. CV for endSonRhymetoH2.

4. Discussion and Conclusion

The analyses of F_0 excursion, rise time, and slope analyses showed a great deal of inter-speaker variability. We found some significant rate effects for each of the three parameters, but no patterns that held for all (or even most) speakers. These results thus fail to overwhelmingly support either the segmental anchoring hypothesis or the constant slope hypothesis. The variability observed is unsurprising since similar results were found in earlier studies of French ([14] and [9]).

We failed to find a stable anchor point for L2. The late rise often began in the onset consonant of the last syllable of the AP or early in the vowel of this syllable; so it might be tempting to suggest that the L2 anchor point is the beginning of the syllable. However, consideration of all the available evidence, from the current study as well as from our previous work and from reports in the literature, suggests that such a conclusion would be premature. The rate differences in L2 alignment found for two speakers are unexpected in a segmental anchoring account. L2 alignment showed considerable variability even for speakers without significant rate effects (see, for example, standard errors for Speakers 3 and 4, Table 2). And finally, there are cases in the current study and in previous studies in which L2 is realized fairly late in the last syllable or in the penultimate syllable.

The lack of constant rise times is evidence against the leading tone hypothesis. Studies for several other languages have also cast doubt on the AM concept of leading and trailing tones linked to associated tones by a fixed temporal interval (see, for example, [3], [15], [16]). Nevertheless, the fact that both L2 and H2 are realized near the end of the AP gives us reason to believe that the two tones form a cohesive unit. No researcher has ever reported an alignment pattern in which L2 is realized toward the middle of the AP rather than toward the end.

The pattern of results found for H2 alignment does not match the predictions of any of our hypotheses. To account for these results, we could postulate different anchor points depending on syllable structure (one for CV and CVCobs, another for CVCson). Yet, apparent differences in alignment based on differences in segmental composition may mask underlying regularities (see, for example, [2], [3]).³ We therefore looked to provide a unified account.

Our account is based on the concept of a segmental “anchorage,” an area where boats can drop anchor. This metaphor more aptly describes the alignment of the peak of the French late rise than the more narrow metaphor of an immovable anchor. Our imaginary anchorage might consist of a stretch of sea bounded by underwater reefs or rocky crags; prudent sailors will drop anchor only where their anchors can sink into the sandy bottom, but never beyond the reefs. In our account, the segmental anchorage for H2 is the region stretching from approximately 20 ms before the end of the vowel to the end of the accentual phrase.⁴ In CV syllables, H2 can thus be realized at the end of the vowel (or just beyond it). In CVCson syllables, H2 can be realized at the end of the vowel or in the coda consonant. In CVCobs syllables, H2 can be realized at the end of the vowel or, in principle, in the voiced part of a voiced obstruent coda. This latter realization is rare, however, possibly due to articulatory constraints. To realize an F_0 peak in a voiced obstruent, the speaker needs not only to maintain vocal fold vibration, but also to increase the rate of vibration. From the point of view of production, simply maintaining voicing during an obstruent requires increased articulatory effort. For the French H2, it may not be worth expending the effort to realize the peak in a voiced obstruent, since the peak can be realized with no extra effort at the end of the vowel and these alignment choices do not differ phonologically. Realizing H2 in the vowel may also preserve

³Many alignment studies, however, have used materials with mostly or exclusively sonorant target sequences ([1], [2], and many others). Thus, for many languages, the influence of syllable structure or segmental composition on alignment is an open question.

⁴We believe that the righthand boundary of the anchorage region is the end of the accentual phrase and not the end of the last full syllable because in cases in which the target AP is produced with a final schwa, the H2 peak can appear in the vowel or coda consonant of the last full syllable, but also in the schwa syllable.

cues to the identity of the voiced coda consonant.

While we use the concept of segmental anchorage to explain peak alignment in the French late rise, we do not claim that it will account for alignment patterns found in all other languages. The segmental anchoring literature demonstrates convincingly that for some languages (e.g., Greek [3]), the start and end points of rises cluster tightly around single anchor points. Indeed, even for French, segmental anchorage may not provide an appropriate account for all alignment patterns. What we do claim, however, is that alignment patterns of the type found for the French H2 will be found for other languages. That is, the segmental anchorage alignment pattern is one of a range of alignment patterns possible in the world’s spoken languages.

5. Acknowledgments

We thank Allison Blodgett, Christian Bulfone, Nino Medves, and Pascal Perrier for their assistance. We are also grateful to our speakers for their participation. This research was supported by a Marie Curie International Fellowship within the 6th European Community Framework Programme.

6. References

- [1] Ladd, D. R., Faulkner, D., Faulkner, H., and Schepman, A. “Constant ‘segmental anchoring’ of F_0 movements under changes in speech rate”, *J. Acoust. Soc. Amer.*, Vol. 106, pp. 1543–1554, 1999.
- [2] Ladd, D., Mennen, I., and Schepman, A. “Phonological conditioning of peak alignment in rising pitch accents in Dutch”, *J. Acoust. Soc. Amer.*, Vol. 107, pp. 2685–2695, 2000.
- [3] Arvaniti, A., Ladd, D. R., and Mennen, I. “Stability of tonal alignment: The case of Greek prenuclear accents”, *J. of Phonetics*, Vol. 26, pp. 3–25, 1998.
- [4] Igarashi, Y. “‘Segmental anchoring’ of F_0 under changes in speech rate: Evidence from Russian”, *Proceedings of the Speech Prosody 2004 Conference*: 25–28, 2004.
- [5] Rossi, M. “L’intonation et l’organisation de l’énoncé”, *Phonetica*, Vol. 42, pp. 135–153, 1985.
- [6] Jun, S.-A. and Fougeron, C. “A phonological model of French intonation”, in *Intonation: Analysis, modelling and technology*, A. Botinis, Ed., Kluwer, Boston, 2000.
- [7] Pierrehumbert, J. B. and Beckman, M. E. *Japanese Tone Structure*, MIT Press, Cambridge, MA, 1988.
- [8] Welby, P. “The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation”, *Doctoral thesis*, The Ohio State University, 2003.
- [9] Welby, P. “French intonational structure: Evidence from tonal alignment”, submitted.
- [10] Post, B. *Tonal and Phrasal Structures in French Intonation*, Thesis, The Hague, 2000.
- [11] Boersma, P. and Weenink, D. “Praat: Doing phonetics by computer” (version 4.3) [computer program]. Retrieved from <http://www.praat.org>, 2005.
- [12] Hermes, D. J. and van Gestel, J. C. “The frequency scale of speech intonation”, *J. Acoust. Soc. Amer.*, Vol. 90, pp. 97–101, 1991.
- [13] ‘t Hart, J., Collier, R., and Cohen, A. *A Perceptual Study of Intonation*, Cambridge University Press, Cambridge, 1990.
- [14] Fougeron, C. and Jun, S.-A. “Rate effects on French intonation: Prosodic organization and phonetic realization”, *J. of Phonetics*, Vol. 26, pp. 45–69, 1998.
- [15] Arvaniti, A., Ladd, D. R., and Mennen, I. “What is a starred tone? Evidence from Greek”, in *Papers in Laboratory Phonology V: Acquisition and the lexicon*, M. B. Broe and J. B. Pierrehumbert, Eds., Cambridge University Press, Cambridge, 2000.
- [16] Grice, M., Ladd, D. R., and Arvaniti, A. “On the place of phrase accents in intonational phonology”, *Phonology*, Vol. 17, no. 2, pp. 143–185, 2000.