Project 5: The Role of Pitch in the Perception of Vlach Stressed Syllables

Introduction

This experiment builds upon my previous project, which described the various factors involved in the production of stress in Vlach. This previous project indicated that stressed syllables in Vlach were distinguished by a combination of factors including vowel quality, intensity, pitch and duration of the syllable. For this project, I performed a perceptual experiment that aimed to determine whether it was possible to shift the perceived stress in a disyllabic word by modifying the pitch alone and controlling for the other factors of vowel quality, intensity, and duration. The hypothesis was that by starting with a sample that was neutral in terms of these three factors, the pitch on each syllable could be adjusted to determine which syllable was perceived to be stressed. However, the experiment failed, with the subject identifying only 15 of 290 tokens as having word-initial stress. To understand why this failure occurred, one must consider two factors: the experimental design and the nature of the sole participant in the experiment. This paper is an attempt to find the weakness in this experiment so that future experiments might be carried out with better success.

Methods

Motivations and Assumptions

The motivation behind this experiment was to quantify the role of various factors involved in the perception of stress in disyllabic words in Vlach. Since it appears that at least four factors (pitch, intensity, duration, and vowel quality) are involved in the production of stress, all four of these factors must be somehow accounted for in the experiment. This experiment focuses on the role of pitch, and it was conceived as if it were to be one part of a four-part experiment: the other three parts of this experiment, if carried out, would focus on the three other factors involved in stress.

A major obstacle lies in the task of measuring one of these factors (or variables) in isolation from the others. One of the assumptions of this experiment, which may be flawed, is that is possible to measure the effects of changes in one variable in isolation from the other variables. The assumption is that there exists (in principle) a token that is neutral in terms of stress for all variables. This neutral token, if it were heard by a native speaker, would be perceived as ambiguous. Furthermore, we might expect that if it were repeated several times and the speaker forced to judge whether it had first-syllable stress or second-syllable stress, the speaker's judgments would be evenly distributed between the two choices.

General Experimental Design

The first step in the design of this experiment, then, was to find a token that was neutral for all stress variables. The second step was to modify this neutral token to create a number of tokens with gradated values of pitch for both the first and second syllables. Finally, the modified tokens, along with several control tokens, were played to a native

speaker in a forced-choice identification experiment in which the participant had to identify whether the stress was on the first or second syllable of the word.

For reasons described below, one word was chosen to be used in the experiment, a combination of the words [ˈɪntɾA] "he enters", and [ɪnˈtɾA] "he entered". The first and second syllables were modified over a range of five pitch values, which combined to produce twenty-five tokens with modified pitch. Additionally, unmodified tokens of [ˈɪntɾA] and [ɪnˈtɾA] were included as controls. Two additional control tokens were included: one was the neutral token with its original pitch contour; the other was the neutral token with a flat pitch contour. These twenty-nine tokens were repeated ten times each to obtain 290 total responses.

Controlling for Vowel Quality

Since it is difficult to modify formant frequencies in Praat, the vowel quality differences between stressed and unstressed syllables could pose a major problem to the design of an experiment such as this one. However, it had been observed in my previous project that the vowel quality does not necessarily change significantly between stressed and unstressed syllables in certain words. Most notably, the final vowel of the word [ˈɪntɾA] was produced either as a low-back vowel or sometimes as a slightly more mid-central vowel, with no apparent effect on the stress. Thus, it was possible to find instances of the minimal pair [ˈɪntɾA] / [ɪnˈtɾA] where the vowel quality was independent of stress.

Below is a chart showing the vowel formants from the first and second syllables of these words, as recorded in the previous project on Vlach stress. While there is a visible difference between the vowel quality of unstressed and stressed vowels (both [ɪ] and [A]), there is overlap between the groups. The arrows indicate the values for two tokens identified as having relatively ambiguous vowel quality: token 09 (an instance of [ɪnˈtɾA]) and token 16 (an instance of [ˈɪntɾA]) from the past project.

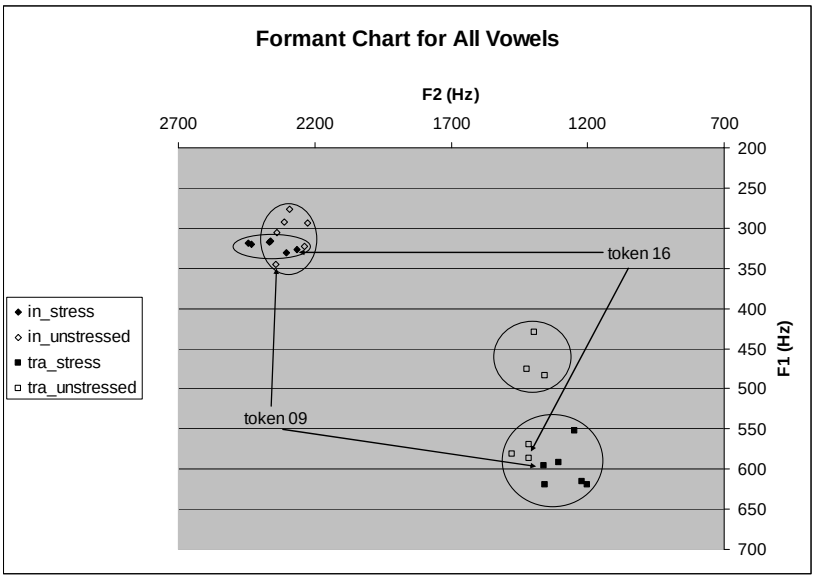

Figure 1: Vowel formants recorded for [ˈɪntɾA] / [ɪnˈtɾA] in project 4.

Due to their neutral vowel quality, these tokens were used as the basis for generating the neutral token that formed the basis of this experiment. (Both token 09 and token 16 were included unmodified as control tokens in this experiment, and both were correctly identified every time. This confirms that they are acceptable examples of both first-syllable and second-syllable stress, despite the lack of any significant vowel reduction. This would seem to justify their use in this experiment.)

Controlling for Intensity

Intensity is a significant factor in stress production (and presumably perception), and in order to create a neutral token, the intensity of the two syllables must be relatively equivalent. However, it turns out that it is very difficult to increase the intensity of an unstressed syllable to the level of a stressed syllable because it is so quiet to begin with that the amplification introduces a significant amount of noise.

Therefore, the decision was made to create an artificial token consisting of the stressed first syllable of token 16 and the stressed second syllable of token 09. This would solve the amplification problem and would not introduce any vowel quality problems, as can be seen in the figure above both tokens have very similar formant values for both syllables. Also, this artificial token should include less inherent bias than a token that was produced as an actual instance of first-syllable or second-syllable stress.

Controlling for Duration

The final variable to control for is duration, which was modified using Praat's duration manipulation feature. The duration values were chosen to be midway between the stressed and unstressed duration for the syllable in question. Thus, using tokens 09 and 16 as references, the first-syllable duration was modified to be equal to 185ms (midway between 160ms and 211ms), and the final syllable was left unmodified at 261ms, as the difference between the two final-syllable durations (278ms and 261ms) was deemed to be statistically negligible.[1]

The duration modifications may have been a weak point in the experiment, as they were based only on the two tokens mentioned above, not on the whole of the data from Project 4 which shows the full range and average values of the durations of these syllables. If the full Project 4 dataset had been used to calculate mean midpoint values for duration, the first-syllable duration would have been 205ms and the second-syllable duration would have been 278ms.

Modifying the Pitch

After visually inspecting the pitch contours of the first-syllable-stressed token and the second-syllable-stressed token, the following design was chosen to modify the pitch contours. The first syllable would have fixed start and endpoints, with a midpoint that would vary over a range of frequencies. The second syllable would have a fixed endpoint and a start point that would vary over a range of frequencies. This design is summarized in the figure below:

[1] Note that the 261ms duration corresponds to the stressed instance of this syllable; normally we would expect the stressed syllable to be longer than the unstressed syllable, but this particular pair is contrary to that general trend. This was a second consideration in not modifying the duration of the syllable.

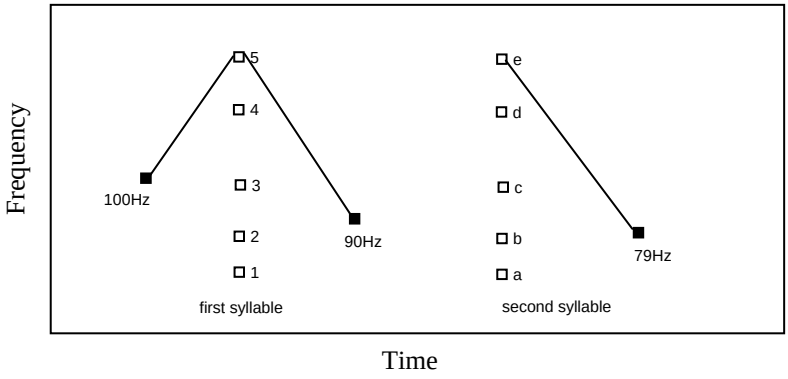

Figure 2: Visualization of the modified pitch contour scheme. The dark squares indicate the fixed pitch points for all tokens, while the clear squares indicate the variable pitch points. The lines indicate the pitch contour of token 5e. The values for pitch points 1-5 and a-e were (75Hz, 85Hz, 100 Hz, 120Hz and 150Hz)

As can be seen from figure 2, the design produced 25 tokens representing the various combinations of each syllables 5 different pitch values. Values 2 and 4 (and b and d) were based on the actual pitch values in tokens 09 and 16, with additional values added below, above, and between these values to create a gradient. Thus, modified token 2d is very similar in pitch to token 09 and modified token 4b is very similar in pitch to token 16. Thus, one would expect to see similar judgments for those pairs of unmodified / modified tokens.

The two figures below demonstrate the similarity of the pitch contours of these tokens:

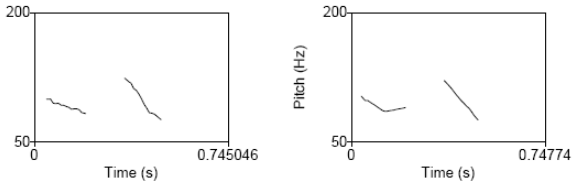

Figure 3: Comparison of pitch contours of tokens 16 (left) and 4b (right)

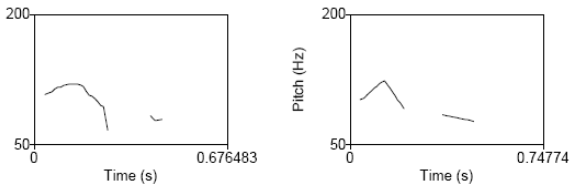

Figure 4: Comparison of pitch contours of tokens 16 (left) and 2d (right)

This design expects token 2d to be identified as first-syllable stress by comparison to token 16, and for token 4b to be identified as second-syllable stress by comparison to token 09. The other tokens between 1a and 5b would be expected to receive judgments based on, perhaps, the relative pitch of the first and second syllables.

Results

The results for this experiment were indicative of some type of experimental failure. Of 290 tokens, only 15 (including all 10 instances of the control token 16) were identified as having first-syllable stress. All others were identified as having second-syllable stress.

Of the remaining 5, these included the first instances of 4b, 5b, 5c and 3b, as well as a late instance of 2a (likely an error). Tokens 4b, 5b, 5c, and 3b are all tokens that would have been hypothesized to be judged as first-syllable stress by the experimental design. Thus, the experiment seemed to be working as planned up to a point. However, once token 16 was played to the listener, he never again identified any token other than 16 (and the late instance of 2a) as first-syllable stress.

Perhaps the most insightful result is that the unmodified base token, which was presumed to be neutral in terms of stress, was consistently identified as having second-syllable stress. This is a strong indicator that the "neutral" token on which all of the modified tokens were based was not, in fact, neutral.

Some of the comments made by the experimental participant may also be insightful. During the experiment he made some comments: first, he asked how he should respond if the word had stress on both syllables; later, he indicated that he had decided to just always choose the second-syllable option in the ambiguous cases. After the experiment had ended, he indicated his suspicion that some of the recordings had been made by splicing a stressed first syllable and a stressed second syllable, although in a later comment he revealed that he was not certain whether this was the case or whether he had in fact produced the doubly-stressed token himself in the previous recording session.

In his critique of the experiment, he stated that although the various tokens were "moving the stress" from the second syllable to the first, they never fully did so in his opinion. This indicates that the pitch alterations did have some effect on his perception of stress, but that the changes were not enough to fully shift the stress to another syllable. Thus, we should not assume that pitch is irrelevant to his perception of stress.

However, it may be that since the tokens were still not fully natural, he sought another indicator of stress on which to base his decision. In his own words, this indicator was the "length" of the final vowel, by which I presume he means duration rather than vowel quality, although I am not entirely sure. This certainly is a result of using the stressed second-syllable in the creation of the neutral token, and it may suggest that the second syllable duration needed to be shortened more. It is unclear why the second-syllable was perceived as more stressed than the first, since both recordings were in the stressed position.

It is also possible that by "length" the participant meant vowel quality. This is a possible interpretation, since the final vowel was of the non-reduced variety. However, the vowel quality was very similar to that of the [ɔintrA] control token, where it did not seem to affect the perception of stress. Nevertheless, it is possible that the participant chose to focus on vowel quality in the ambiguous, modified tokens, and that in token 16 the vowel quality could be neglected due to the presence of other indicators of stress.

The participant also had questions about whether the experiment recorded his response time to answer the question. This, along with his other observations, indicates that he was actively trying to see through the purpose and design of the experiment, and the results must be considered at least somewhat distorted as a result of this.

Conclusions

To be entirely honest, this experiment may have had design flaws that arose from its hasty preparation. The experiment could have been more robust by including modifications of other factors than pitch, but this was considered too complicated for one

project. The flaws of this experiment could be corrected for in future experiments, but this would require more time and probably more than one speaker.

A naturalness judgment option was not included in this experiment, but one might have been helpful. From the participant's comments, it seems likely that all of the modified tokens would have been identified as unnatural to some extent.

One problem with the experiment might be that the intensity of both syllables was left unchanged. While this was done based on the assumption that it is only relative intensity, and not absolute intensity, that matters for stress perception, it would have been a much better idea to reduce the intensity of both syllables to a more neutral value. It may be that the high intensity of the second syllable is what cued the listener to place the stress on the second syllable. Although the first syllable had similar intensity, the intensity may have been perceived as higher for the second syllable by virtue of the fact that the latter part of a word will tend to have lower intensity than the first when no other factors are present.

One conclusion that can be drawn about the perception of stress in Vlach is that there are many factors involved, and modifying only one of these factors results in an unclear stress pattern. In the absence of other factors (i.e. if the control tokens had been absent) it may be that stress alone is enough to shift the stress of a word. This seems to be supported by the very early responses in the experiment, but cannot be confirmed due to the disturbed results for the majority of the experiment.

While one might conclude that this experiment shows that pitch is not a significant factor in the perception of stress in Vlach, I am not too quick to agree with that conclusion. In my opinion, the experiment was flawed and I would like to see the results of using these same tokens, without (or in a separate session from) the control tokens, and using Vlach speakers other than the one in this experiment.

Appendix

Target Sentence

Each token was surrounded by a target sentence, which was the same as in the previous project:
[αY σπυν λ↔καρδιA _____ λ ℘ ϖλ↔Σ↔Στ↔]
"I say the word _____ in Vlach."

Experiment Responses

In the list of responses below, the sequence number, prompted token, and response are listed. A response of 1 indicates first-syllable stress, a response of 2 second-syllable stress. The stimulus "unmod" represents the neutral token with its original pitch contour, and "flat" is the neutral token with a flat pitch contour.

number	stimulus	response	35	2b	2	70	3c	2
1	5d	2	36	2d	2	71	4e	2
2	4b	1	37	4e	2	72	2c	2
3	5b	1	38	flat	2	73	4a	2
4	3e	2	39	control09	2	74	4d	2
5	4e	2	40	2e	2	75	3e	2
6	3a	2	41	5c	2	76	6e	2
7	2c	2	42	6c	2	77	control16	1
8	5c	1	43	6e	2	78	6c	2
9	2a	2	44	unmod	2	79	6d	2
10	4d	2	45	control16	1	80	6a	2
11	4c	2	46	4c	2	81	2b	2
12	5e	2	47	5b	2	82	flat	2
13	control09	2	48	6d	2	83	2d	2
14	3b	1	49	4a	2	84	4c	2
15	6b	2	50	6a	2	85	6b	2
16	6d	2	51	3b	2	86	5b	2
17	6e	2	52	4b	2	87	2a	2
18	3c	2	53	5e	2	88	unmod	2
19	2e	2	54	6b	2	89	control09	2
20	4a	2	55	3e	2	90	6e	2
21	unmod	2	56	2a	2	91	6a	2
22	control16	1	57	4d	2	92	6b	2
23	6c	2	58	3a	2	93	5a	2
24	2d	2	59	control09	2	94	2d	2
25	2b	2	60	3b	2	95	3e	2
26	5a	2	61	4b	2	96	5c	2
27	6a	2	62	5c	2	97	2c	2
28	flat	2	63	unmod	2	98	5e	2
29	3d	2	64	5a	2	99	4c	2
30	3c	2	65	2e	2	100	6d	2
31	2c	2	66	5d	2	101	5d	2
32	5a	2	67	3d	2	102	4a	2
33	5d	2	68	5e	2	103	3a	2
34	3d	2	69	3a	2	104	flat	2

105	2e	2	155	2d	2	205	6c	2
106	2a	2	156	3a	2	206	4c	2
107	2b	2	157	5d	2	207	5e	2
108	4e	2	158	4a	2	208	3c	2
109	4b	2	159	2b	2	209	5c	2
110	3c	2	160	6d	2	210	unmod	2
111	5b	2	161	5c	2	211	control16	1
112	3d	2	162	4c	2	212	3e	2
113	6c	2	163	2a	2	213	3d	2
114	4d	2	164	unmod	2	214	2b	2
115	3b	2	165	3d	2	215	5a	2
116	control16	1	166	2c	2	216	control09	2
117	flat	2	167	5a	2	217	6e	2
118	5a	2	168	2e	2	218	flat	2
119	4d	2	169	6b	2	219	6b	2
120	control09	2	170	3c	2	220	4e	2
121	5c	2	171	4b	2	221	2e	2
122	4c	2	172	control16	1	222	4a	2
123	2c	2	173	flat	2	223	3a	2
124	4e	2	174	4d	2	224	5d	2
125	3c	2	175	6c	2	225	6a	2
126	2e	2	176	3c	2	226	3b	2
127	4b	2	177	3b	2	227	2d	2
128	4a	2	178	4e	2	228	2c	2
129	6d	2	179	flat	2	229	4d	2
130	2d	2	180	4b	2	230	2a	2
131	6c	2	181	6d	2	231	5b	2
132	6e	2	182	control09	2	232	4b	2
133	control16	1	183	2e	2	233	6e	2
134	unmod	2	184	5a	2	234	3e	2
135	5b	2	185	2c	2	235	4d	2
136	2a	2	186	4a	2	236	5b	2
137	3b	2	187	2d	2	237	3c	2
138	5d	2	188	2a	2	238	2e	2
139	2b	2	189	5b	2	239	4c	2
140	5e	2	190	2b	2	240	4a	2
141	3a	2	191	6a	2	241	3d	2
142	3e	2	192	6b	2	242	6b	2
143	3d	2	193	4c	2	243	2c	2
144	6a	2	194	5e	2	244	5d	2
145	6b	2	195	4d	2	245	flat	2
146	5b	2	196	5d	2	246	3b	2
147	control09	2	197	control16	1	247	control09	2
148	3e	2	198	5c	2	248	4e	2
149	5e	2	199	3d	2	249	6d	2
150	4e	2	200	unmod	2	250	unmod	2
151	6a	2	201	6e	2	251	control16	1
152	6e	2	202	3e	2	252	2a	1
153	6c	2	203	3a	2	253	5a	2
154	3b	2	204	6d	2	254	5e	2

255	6c	2
256	4b	2
257	6a	2
258	2d	2
259	3a	2
260	2b	2
261	5c	2
262	5b	2
263	3a	2
264	control16	1
265	2e	2
266	3c	2
267	4c	2
268	2b	2
269	4b	2
270	flat	2
271	6a	2
272	3e	2
273	5c	2
274	6c	2
275	5e	2
276	6b	2
277	control09	2
278	4a	2
279	5a	2
280	5d	2
281	2d	2
282	6e	2
283	unmod	2
284	4e	2
285	6d	2
286	3b	2
287	2c	2
288	3d	2
289	4d	2
290	2a	2