UNIVERSITY OF ALBERTA

# L1 AND L2 ACCENT ALIGNMENT IN GERMAN AND CANADIAN ENGLISH

BY

KRISTIN E. SUDU

A THESIS
SUBMITTED TO THE FACULTY OF ARTS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF ARTS

DEPARTMENT OF LINGUISTICS
DEPARTMENT OF MODERN LANGUAGES AND CULTURAL STUDIES

EDMONTON, ALBERTA
April, 2019

UNIVERSITY OF ALBERTA
FACULTY OF ARTS

The undersigned certify that they have read and recommend to the Faculty of Arts for acceptance, a thesis entitled *L1 and L2 Accent Alignment in German and Canadian English*, submitted by Kristin Sudu in partial fulfillment of the requirements for the degree of Bachelor of Arts.

…….

Honors Thesis Supervisor

…….

For the Department

…….

Honors Thesis Supervisor

…….

For the Department

# ABSTRACT

This study is a replication of a previous study done by Atterer & Ladd (2004) looking at the prosodic contours of English and German bilinguals. Atterer & Ladd (2004) showed that segmental anchoring of prenuclear rising accents occurs in both English and German, but speakers of these languages align the L(ow) and H(igh) tones of these accents at different points within the segmental material. Native alignment patterns were shown to be transferred into their L2 when German speakers spoke English. In this study I show that native Canadian English speakers also transfer native alignment patterns into their L2 (German). I also found that Canadian English speakers' alignment patterns were significantly later than those of the British English control group used by Atterer & Ladd (2004).

ACKNOWLEDGEMENTS

I must thank my amazing, supportive supervisors who were there for me every step of the way. To Dr. D, who changed my life forever by seeing potential in me that I could not even see myself. And to Anja, who always had the perfect way of explaining things and helping me through my many roadblocks. You both have taught me so much and I am forever grateful.

Also, many thanks to my friends and family who supported me throughout the thesis designing, conducting, and writing process; your limitless amount of patience was more appreciated than you know.

TABLE OF CONTENTS

TABLES AND FIGURES

## SECTION 1 – BACKGROUND INFORMATION

The theoretical underpinnings of this paper had their beginnings in the early 1970's, when researchers began to find ways of connecting "instrumentally recorded $f_0$" – intonation – with phonological structure, a concept later called "intonational phonology" (Ladd, 2008). During this time two major approaches were developed, sparking a debate known as "levels vs. configurations" (Arvaniti, Ladd, & Mennen, 1998). The supporters of the "configurations" interpretation posit that "the basic entity of intonation is a pattern", i.e. the configuration of $f_0$ contours is created by the movements between $f_0$ peaks and valleys (Bolinger, 1951). However, this paper will examine current theories that developed from the "levels" argument. The first founding study of intonational phonology as we know it today is Bruce (1977). He studied variations in f0 contours caused by word accent and focus marking in Stockholm and Malmö dialects of Swedish. His contribution was to establish that the word accent distinction between these two dialects was created by a consistent difference in alignment of the $f_0$ peak, and not a difference in the shape of the $f_0$ contour; i.e. the $f_0$ movement itself is unimportant, what matters is that precise 'tonal targets' are reached at the correct time within the segmental material (Atterer & Ladd, 2004; Arvaniti, Ladd, & Mennen, 1998).

Another founding study for the field of intonational phonology is Liberman (1975), in which he conceives his "Metrical theory" to describe English intonation. He posits "tune" (intonation) as a completely independent entity from "text" (segmental material) and created a two-part system to explain how they align (Liberman, 1975). Firstly, metrical patterns impose structure on complex events and secondly, metrical grids impose structure on the dimension of time (Liberman, 1975). Metrical patterns are "trees with uniformly binary branching, and a relation strong/weak defined on the two elements of each nonterminal constituent" (Liberman, 1975). Metrical grids are "an

ordered set of ordered sets", basically a tree structure in grid form that allows us to relate the intonational structure to time (Liberman, 1975). Shortly after Liberman published his Metrical theory, another "Autosegmental theory" was produced in Goldsmith (1976). Goldsmith used his studies of African lexical tones along with an analysis of intonation in English, to propose a non-linear representation of "autosegments" (tones and phonological segments) which were placed into syllabic and tonal tiers and associated together by "association lines" (Goldsmith, 1976). From this came his "starred notation" where the tone associated with the accented syllable of a word is denoted with an asterisk (Goldsmith, 1976). This starred notation was adapted for annotation alignment distinctions by Pierrehumbert in her 1980 dissertation *The phonology and phonetics of English intonation* (Atterer & Ladd, 2004). She proposed that the bitonal tones of word accents have a "starred" tone, which is aligned with the accented syllable and an "unstarred" or "trailing" tone, which she marks with a raised hyphen (Pierrehumbert, 1980). She compares these tones to the metrical feet proposed by Liberman (1975): the starred tone is similar to the strong syllable of a foot, and the trailing tone is similar to the weak syllable of a foot (Pierrehumbert, 1980).

The ideas put forth by these papers have been synthesized into the current "Autosegmental-Metrical (AM) theory" (Ladd, 2008). The most important aspects of AM theory are that "[1.] the primitives of intonation systems are $f_0$ targets or level tones, and that [2.] $f_0$ movements are only transitions or interpolations between tonal targets" (Arvaniti, Ladd, & Mennen, 1998). These level tones or "pitch targets" comprise pitch accents and boundary tones in intonational languages and can be either High (H) or Low (L) (Ladd, 2008). Within AM theory there is the notion of "segmental anchoring" first touched upon by Arvaniti, Ladd, & Mennen (1998) in their study of prenuclear accents in modern Greek. They found that each of the tones in the *L+H bitonal pitch accent had its own "anchoring point" within the segmental string (Arvaniti, Ladd, & Mennen,

1998). They determined this by testing the relationship between the duration of the $f_0$ excursion ("the 'horizontal' time dimension") and the change in pitch ("the 'vertical' $f_0$ dimension") (Arvaniti, Ladd, & Mennen, 1998). They refer to the phonological and phonetic properties corresponding to each of these dimensions as *alignment* and *scaling* (Arvaniti, Ladd, & Mennen, 1998). They found no correlation between the length of the $f_0$ pitch excursion and the change in pitch, i.e there was no fixed "distance" between the two tonal targets (Arvaniti, Ladd, & Mennen, 1998). This led them to conclude that tonal targets are indeed anchored to segmental targets (Arvaniti, Ladd, & Mennen, 1998). Similar results were found for British English by Ladd, Faulkner, Faulkner, & Schepman (1999). They tested the segmental anchoring hypothesis against more traditional "'constant slope' and 'constant duration' hypotheses" by comparing $f_0$ contours and tonal alignment at various speech rates (slow, normal, and fast). They hoped to prove that slope and duration are not characteristic to particular accent types and that accent shape references the $f_0$ level and alignment of the accent's targets (Ladd et al., 1999). Their data showed that speech rate did have a significant effect on alignment ($f_0$ duration, specifically rise time) but not on scaling (f0 excursion), proving both the constant slope and constant duration hypotheses inadequate (Ladd et al., 1999). Supporting these findings is a study by Prieto, van Santen, & Hirschberg (1995), looking at $f_0$ peak alignment in Spanish. Like Ladd et al. (1999). they found that any peak delay effects had the greatest impact on the upslope of a rise, or the alignment of the H tone (Prieto et al., 1995). The beginning of the rise (the L tone) was shown to be very stably anchored within the onset/rhyme of the accented syllable (Prieto et al., 1995). This study also showed differing effects of vowel or onset lengthening on peak delay: when lengthening was caused by slower speech rates $f_0$ peaks were also delayed, but when lengthening was due to upcoming prosodic events (e.g. phrase

boundaries) peaks were not delayed (Prieto et al., 1995). This difference in delay effect was also obtained in the results of Silverman & Pierrehumbert (1980).

Finally, I discuss the results of the replicated study Atterer & Ladd (2004). Researchers conducted two experiments analysing segmental anchoring of $f_0$ in two different German dialects and in the L2 English of native German speakers. The authors aimed to compare these results to previous findings regarding prenuclear accent alignment in Greek and Dutch. The first experiment compared the prenuclear rising (L*+H) accents of German speakers who either spoke a Northern or Southern dialect of Standard German. 18 participants were asked to read prepared test sentences aloud, where segmental composition and phrase position of test words were controlled. Their results were as follows:

> "…although the mean overall duration of the CVC sequence is virtually identical in the two groups of speakers, the proportion of the vowel duration is greater in the Southern speakers than in the Northern speakers. Specifically, in Southern speech the consonant preceding the stressed vowel is relatively shorter and the stressed vowel is relatively longer, compared to Northern speakers" (Atterer & Ladd, 2004).

In the second experiment they had the same native German speakers read prepared English test sentences originally used in Ladd et al. (1999) where they tested the effect of speech rate on segmental anchoring. This allowed for direct comparison of the L2 speakers and the native English results observed in Ladd et al. (1999). The speakers were described as not "near-native" in their proficiency and 15 were described as having readily identifiable German or "markedly foreign" accents in English (Atterer & Ladd, 2004). The authors predicted that their native German alignment patterns would be transferred into their L2 English (Atterer & Ladd, 2004). They observed a consistently later alignment of the L in German speakers compared to native English speakers, as well as a probable later alignment of the H (Atterer & Ladd, 2004). They also found

that the native alignment pattern of their dialects was carried over into their English, i.e. Southern dialect speakers still aligned later than Northern dialect speakers in English as well as in German (Atterer & Ladd, 2004).

Their predictions and results concerning crosslinguistic transfer of alignment patterns are supported by previous findings from Jilka (2000), who studied English speakers with advanced competency in German, and Mennen (2004) who looked at Dutch speakers with near-native fluency in Greek. It was found in these studies that adult-acquired second languages are vulnerable to transfer of native alignment patterns (Atterer & Ladd, 2004). Jilka (2000) also conducted two experiments which sought to define "intonational foreign accent". The first was a production experiment where a total of 20 bilingual speakers of English and German (10 native German speakers and 10 native English speakers) were asked to read prepared materials aloud and participated in spontaneous conversation. The second experiment was an online perception test completed by native speakers of both languages. From the results of these experiments he derived four major types of intonational foreign accent, including "transfer of tonal category from the L1" (Jilka, 2000). He observed that "native intonation categories may replace their L2 counterparts in an equivalent environment" and compares it to "the transfer of a phonological category on the segmental level" (Jilka, 2000). More recent research conducted by Graham & Post (2017) examined the influence of native language typology and L2 proficiency on the transfer of L1 intonation contours and alignment. The effect of typological similarity between L1 and L2 was tested by comparing native American English speakers to native speakers of two different languages: Puerto Rican Spanish (PRS), a typologically similar language, and Tokyo Japanese, a typologically different language (Graham & Post, 2017). Speakers of varying experience levels were chosen in each group to see if greater L2 proficiency would lead to more native-like

intonation (Graham & Post, 2017). PRS speakers used more target-like contours than the Japanese speakers but Japanese speakers showed greater improvement towards target patterns with experience (Graham & Post, 2017). L1 background and L2 proficiency had no significant effects on tonal alignment, which the authors attributed to the sensitivity of tonal alignment to underlying phonological representations which would be difficult for late L2 learners to acquire (Graham & Post, 2017).

As my research is an almost direct replication of Atterer & Ladd (2004), it seeks to answer very similar questions as the original study:

1) Do Canadian English and/or German speakers align tonal targets to segmental landmarks? If so, to which segmental landmarks are they aligned?
2) Do Canadian English and/or German speakers exhibit cross-linguistic transfer of alignment patterns?

I hypothesize that:

1) Canadian English speakers will have earlier alignment patterns, with tonal targets aligned to similar segmental landmarks as found in Atterer & Ladd (2004). German speakers will have later alignment patterns, with tonal targets aligned to similar segmental landmarks found in Atterer & Ladd (2004).
2) There will be transfer of alignment patterns. Canadian English speakers will have early alignment patterns typical to English when speaking in English and German. German speakers will have later alignment patterns typical of Southern German speakers (Atterer & Ladd, 2004) when speaking in English and German. (I hypothesize Southern German alignment patterns specifically for reasons explained in §2.1).

I expect no significant effect of proficiency among the late-acquired L2 speakers, similar to Graham &Post (2017). My study differs from previous research is that my L2 speakers are native speakers of Canadian English. Other studies have only looked at British or American dialects of English.

## SECTION 2 – METHODS

*2.1 Participants*: All participants were adults organized into a control and an experimental group. Originally, 10 participants were recorded for each group, but due to time constraints only the first five participants to be recorded from each group were chosen for analysis. Two exceptions were made to this: firstly, the sixth English speaker's data was used in lieu of the second's which was unusable due to disfluencies. Secondly, because the English speaker group contained a male participant, the fifth German speaker was replaced by the sole male German speaker to keep the gender composition in each group consistent. Ethics approval was obtained from the University of Alberta (Pro#00081792).

The control group consisted of five native German speakers recorded at Goethe-University Frankfurt am Main, Germany. My supervisor Dr. Arnhold recruited and recorded these participants on my behalf while teaching and conducting her own research at the Goethe-University Frankfurt am Main. They were student participants aged 25-32 (mean = 28) with 4 female and 1 male participant. All speakers came from regions where Middle or Southern dialects of German are spoken. The speakers with Middle German dialects created a problem of classification, since they needed to be coded as either "Northern" or "Southern" German speakers to make my own data and the data from Atterer & Ladd (2004) comparable. Ultimately, they were classified as Southern German speakers, since they come from areas below the Benrather Line – the largest and arguably most important bundle of isoglosses in German dialectology – that separates Northern German from Middle and Southern German (König, 1994). Their years of English experience ranged from 17-20 years (mean = 19.6). They were compensated with €5.00 for every 30 minutes of experiment time.

The experimental group consisted of five students from the University of Alberta. They were aged 18-21 with 4 female and 1 male speakers. Their years of German experience ranged from 2-15 (mean = 6.6). Students received course credit for their participation.

*2.2 Materials*: For this experiment I used the same test sentences as Atterer & Ladd (2004) with some adjustments and additions made. The phenomenon of interest was the rising $f_0$ pitch accent associated with stressed syllables in German and English. I looked at the stressed syllable of the first content word in 30 sentences – 15 in each language. These test syllables met specific criteria in both English and German to make the data comparable across the two languages and to facilitate identification and annotation of $f_0$ minima and maxima (Atterer & Ladd, 2004). Their criteria were as follows:

- "The test word was either an adjective followed by a noun or a noun followed by a genitive construction. This normally ensured that a prenuclear rising accent was put on the test word followed by a nuclear accent on the following noun.
- The test syllable was preceded by one or two unstressed syllables, and followed by one or more unstressed syllables.
- The test syllable always contained a phonologically short (lax) vowel, in order to avoid any effect of vowel length of the sort found in Dutch by Ladd et al. (2000).
- The consonants of the test syllable were always sonorants (normally nasals, occasionally /l/), to ensure a smooth F0 contour." (Atterer & Ladd, 2004: 182)

Two example sentences – one from each language with the test syllables in bold – are given below, but tables of the altered sentences and a full list of the speech materials used are available in Appendix A.

German: Die Ver**läng**erung der Ausleihfrist ist leider nicht möglich.

"The extension of the loan period is unfortunately not possible."

English: There was a **nom**inal fee for his services.

Two German test sentences were shortened so that they better matched the sentence length of the other items and to make them easier for the L2 speakers to understand. I deemed this necessary, since the L2 speakers I would be recruiting were assumed to be second- or third-year students whose German proficiency levels may not have been high enough to fully understand the vocabulary used in the test sentences. I also added two more German sentences to make the number of German and English sentences equal. I changed one word in an English sentence that was deemed confusing. I then made our own filler sentences - 15 for each language - as well as two practice German sentences. The filler sentences were structured similarly to the test sentences but did not contain test syllables that met the required criteria. The German practice sentences also lacked test syllables and were used to prepare the L2 German speakers for the level of German they would be reading.

*2.3 Recording and annotation procedures*: Each speaker group was asked to read both the English and German sentences aloud under laboratory recording conditions in a single session. The order of which language they read first was counterbalanced between subjects. For the native German speakers, the materials were printed onto individual pieces of paper and randomly shuffled for each session. For the L2 speakers the materials were displayed individually on a monitor in random

order. No explicit instructions were given as to how to read the sentences or which words to emphasize.

Test trials were extracted from the recordings and annotated using the computer program Praat (Boersma & Weenink, 2019), which I used to label segmental landmarks and find $f_0$ values. As done by Atterer & Ladd (2004), the following six landmarks were identified using a Praat script for each utterance:

- "C0—the beginning of the initial consonant of the test syllable;
- V0—the beginning of the vowel of the test syllable;
- C1—the beginning of the final consonant of the test syllable;
- V1—the end of the final consonant of the test syllable, i.e., the beginning of the vowel of the following syllable;
- L—the beginning of the F0 rise (local F0 minimum); and
- H—the end of the F0 rise (local F0 maximum)." (Atterer & Ladd, 2004: 183-184)

The Praat script I used automatically created annotation tiers but boundaries for word, syllable, and segment were manually placed. The $f_0$ minima and maxima (L and H) within the target word were automatically identified by the Praat script, but in many cases, they had to be manually changed. This was because the lowest or highest frequency point in the target word was not always the peaks of the accent for the test syllable. For example, in some instances there was a secondary peak after the accent of the test syllable that had a higher $f_0$ frequency. I determined these peaks to be the "elbows" of the greatest change in slope. In total, I recorded 300 test trials: 15 test trials in each language - a total of 30 trials per participant - for 10 participants. However, 79 items of the 300 recorded were discarded due to strange recording quality, an incorrect or mispronounced target

word, different accent on the target word, disfluencies, or other. Numbers and relative percentages of discarded data can be found in Table 1 below.

Discarded Trials

| | Strange Recording Quality | Wrong/Mispronounced Target Word | Different Accent | Disfluencies | Other |
|---|---|---|---|---|---|
| Sorted Out | 13 | 17 | 16 | 31 | 2 |
| Percent of Sorted Out | 16% | 22% | 20% | 39% | 3% |

Table 1. Sorted out trials placed into categories indicating reason for being discarded. First row contains raw numbers of trials sorted out; second row contains the percentage that column constitutes of the total discarded trials.

## Section 3 – Results

My results are divided into four subsections. The first subsection discusses (test) word and segment duration, since it was found to have a significant effect in some of my alignment data. It is further divided into separate sections for each 'type' of duration. In the second subsection I present findings for the scaling, or f0 height, of the L and H tones. The third subsection provides answers to if and/or how speakers align tonal targets to segmental landmarks. Lastly, the fourth subsection presents findings for the effects of native speaker and trial language on alignment of L and H. It is further divided by tonal target-segmental landmark pair for clarity's sake.

*3.1 Differences in word and segment duration:* First, I conducted a series of two-way ANOVAs with each 'type' of duration, i.e. total word, total syllable, onset, etc., as the dependent variable and NATIVE SPEAKER LANGUAGE and TRIAL LANGUAGE as the independent variables. These were completed for each speaker and trial language combination using the mean values found in Table 2. I also completed post-hoc tests for each ANOVA (Tukey HSD) to determine the sources of any main or interaction effects. In sections 3.1.1-5, reported p-values are taken from these post-hoc tests. Next, I calculated the number of syllables and segments per word for the English and the German test words. English test words had a mean of 3.4 syllables/word and 7.2 segments/word, whereas German test words had a mean of 3.53 syllables/word and 8.8 segments/word. These measures were used to explain the longer mean total word duration times of German speakers in German trials, since they would be expected to have shorter total word duration times in their native language compared to their L2.

*3.1.1 Differences in total word duration across native speaker and trial language:* The higher amount of syllables and segments per word for German trials suggests that all participants should have longer total word durations for German trials compared to English trials, regardless of native speaker language. In post-hoc tests, this was confirmed only for the English speaker group, who did have significantly longer total word durations in German trials compared to English trials ($p < 0.01$). However, there was no such significant difference in total word duration across trial languages within the German speaker group. The difference in total word duration for German trials, where English speakers had significantly longer total word durations when speaking German compared to German speakers speaking German, was also shown to be significant in post-hoc tests ($p < 0.01$).

*3.1.2 Differences in syllable duration across native speaker and trial language:* The only significant difference found for syllable duration was that English speakers have longer syllable durations than German speakers for both trial languages ($p < 0.05$).

*3.1.3 Differences in onset duration across native speaker and trial language:* There were no significant effects of either native speaker or trial language found for onset duration.

*3.1.4 Differences in nucleus duration across native speaker and trial language:* The only significant difference found for nucleus duration was that English speakers have longer nucleus durations than German speakers for German trials ($p < 0.05$).

*3.1.5 Differences in coda duration across native speaker and trial language:* The only significant difference found for coda duration was that English speakers have longer coda durations for German trials compared to English trials ($p < 0.05$).

Mean Word and Segment Duration Data by Speaker Group and Trial Language

| Speaker Group & Trial Language (TL) | Total Word Duration | Total Syllable Duration | Onset Duration (% total) | Nucleus Duration (% total) | Coda Duration (% total) |
|---|---|---|---|---|---|
| ***English*** | 675.85 | 250.10 | 105.32 (42%) | 92.91 (37%) | 51.87 (21%) |
| *TL: English* | 507.17 | 221.63 | 90.75 | 87.23 | 43.65 |
| *TL: German* | 898.20 | 287.63 | 124.51 | 100.40 | 62.71 |
| ***German*** | 552.77 | 200.94 | 78.16 (39%) | 69.18 (34%) | 53.60 (27%) |
| *TL: English* | 542.09 | 205.12 | 90.08 | 65.10 | 49.95 |
| *TL: German* | 561.09 | 197.68 | 68.87 | 72.37 | 56.44 |

Table 2. The columns show the mean duration in ms for the test word and each of the segments of the test syllable. Rows show overall durations, English trial durations, and German trial durations for each speaker group. The percentage values indicate the proportion of the test syllable dedicated to that segment.

*3.2 Tonal target $f_0$ height:* I tested for any possible differences in $f_0$ height for either of the tonal targets using the same ANOVA procedure as described in §3.1, but with $F_0$ as the dependent variable. The ANOVAs for L and H were completed separately using the mean values found in Table 4. I found no significant difference ($p > 0.05$) across speaker groups for the $f_0$ of either L or H tones, permitting me to conclude that there is no difference in scaling between speaker groups.

Mean Frequency for L and H Tones by Speaker Group and Trial Language

| Speaker Group & Trial Language (TL) | L Tone Frequency (Hz) | H Tone Frequency (Hz) |
|---|---|---|
| ***English*** | 196.10 | 255.06 |
| *TL: English* | 191.27 | 237.57 |
| *TL: German* | 202.46 | 278.12 |
| | | |
| ***German*** | 187.38 | 264.91 |
| *TL: English* | 190.56 | 246.27 |
| *TL: German* | 184.90 | 279.44 |

Table 3. Mean $f_0$ frequency values in Hz for each speaker group, further divided by trial language.

*3.3 Alignment of L to C0 or V0:* To determine whether the L tone is aligned with C0 (onset beginning) or V0 (nucleus beginning) I conducted various paired t-tests across speaker groups and trial languages. The values I used for my tests were the mean distances of L from either C0 or V0 for each speaker (hereby referred to as L1 and L2 respectively). I started by conducting two-tailed paired t-tests comparing L1 and L2 and found that there was a highly significant difference for all speakers in both languages between the two measures ($p < 0.0001$). I conducted one-tailed paired t-tests which showed that L2 was significantly smaller for all speakers in both languages ($p < 0.0001$), indicating that speakers align their L tones with V0 rather than C0 when speaking either English or German, since smaller values mean the L tone is closer to the landmark. This result was slightly more significant for the German speakers ($p = 1.232e-08$) than the English speakers ($p = 2.059e-05$), however, since both p-values are so small it is unlikely there exists a real significant difference between the speaker groups. A visual representation of the alignment patterns for each speaker and trial language group can be found in Figure 1. For this figure I relativized the L and H tones to the beginning of the syllable by subtracting the time the syllable began from the time which L and H occurred respectively. When looking at Figure 1, it is clear that no L values are actually aligned near the V0 boundary, but rather in the middle of C0 (German speaker L tones)

and V0 (English speaker L tones). What the t-tests tell us then is that the German speaker L tones, while still aligned in the middle of the C0 segment, are slightly closer to the V0 landmark than the C0 landmark. For the English speaker L tones, the t-tests just explain what is visually obvious, that they are much closer to being aligned with V0 than with C0. These t-test results are also supported by the mean alignment values in Table 3 which are much smaller for L2 than for L1, especially for the English speakers.

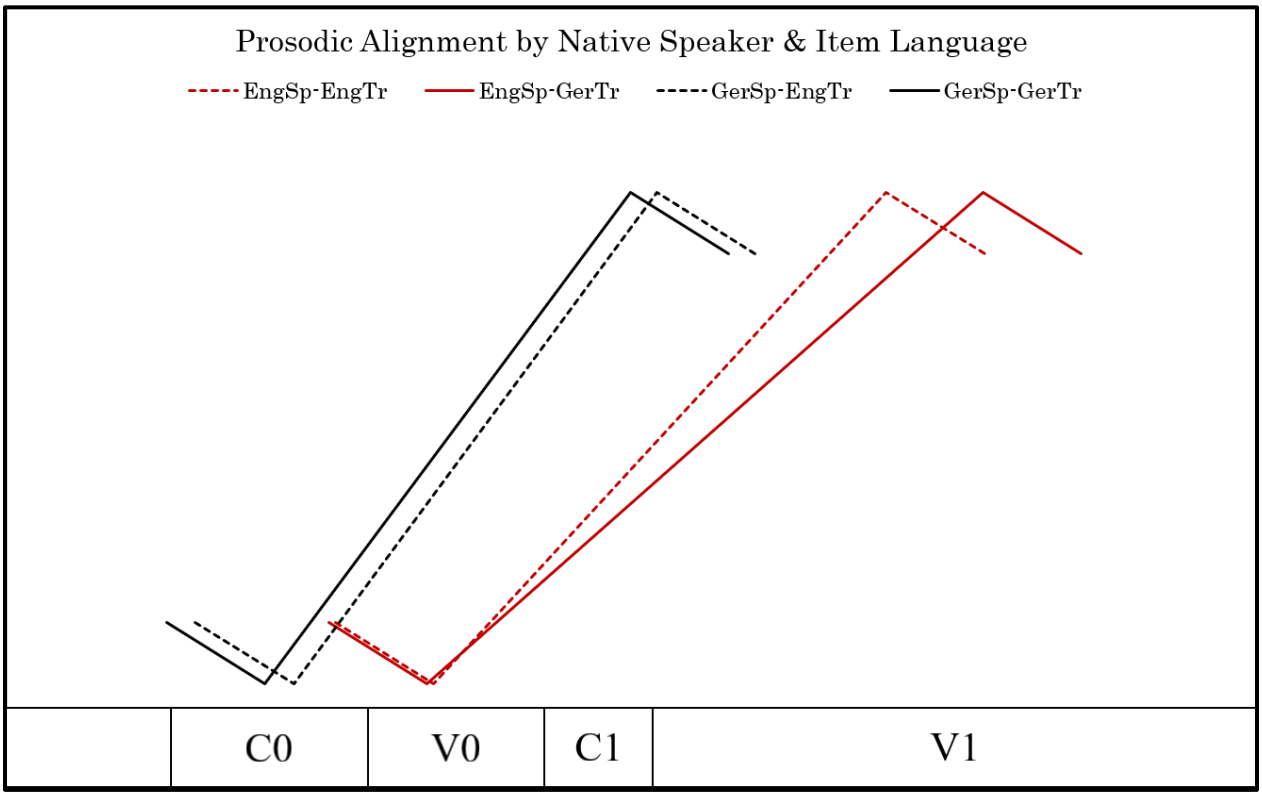


Fig. 1. Schematic representation of the alignment of rises relative to a stressed syllable for German and English speakers divided by trial language. Duration times for the stressed syllable segments (C0, V0, C1) were calculated from grand mean of all speakers for each segment.

*3.4 Independent variable effects on alignment:* Table 3 contains mean alignment values, i.e. mean distance of the tonal target from the segmental landmark, for each of the speakers as well as the grand means for the English and German speakers. Distance of the tonal target from the segmental landmark was calculated by subtracting the time (in ms) of the segmental landmark from the time

(in ms) of the tonal target. I used these mean values to conduct separate two-way ANOVAs for each tonal target-segmental landmark pair, using TONE DISTANCE FROM LANDMARK (L1, L2, or H) as the dependent variable and NATIVE SPEAKER LANGUAGE and TRIAL LANGUAGE as my independent variables. After that I completed post-hoc tests (mainly Tukey HSD in combination with visual interpretation of interaction plots) to determine the sources of any main or interaction effects.

Mean Alignment Data

| Speaker | L1 (ref C0) | L2 (ref V0) | H (ref V1) |
|---|---|---|---|
| *English speakers* | | | |
| E01 | 63.72 | -16.73 | 100.57 |
| E03 | 188.39 | 29.21 | 138.23 |
| E04 | 108.47 | -12.57 | 30.81 |
| E05 | 134.36 | 57.23 | 105.88 |
| E07 | 111.46 | 20.53 | 103.56 |
| English grand mean | 118.76 | 13.45 | 95.80 |
| *German speakers* | | | |
| G01 | 52.54 | -24.34 | 5.16 |
| G02 | 44.96 | -31.32 | 22.80 |
| G03 | 54.64 | -15.36 | 38.95 |
| G04 | 64.19 | -22.90 | 21.88 |
| G06 | 30.38 | -50.51 | -13.37 |
| German grand mean | 49.11 | -29.05 | 14.40 |

Table 4. The columns show the distance in ms between a tonal target (L or H) and a segmental landmark (C0, V0, or V1). A negative value indicates that the $f_0$ label occurs before the segmental label.

*3.4.1. Effect of native speaker language and trial language on L alignment to C0:* When conducting my two-way ANOVA for L1, I had to include word duration as a third variable. This was because the interaction effect between speaker language and trial language would disappear if word duration was removed from my model, even though word duration did not have any interaction effects with either native speaker language or trial language. The main effect of word

duration had a significance level of $p < 0.001$. The interaction effect between trial language and native speaker language had a significance level of $p < 0.01$. Post-hoc tests revealed a significant difference between the English speakers speaking English and all other native speaker/trial language groups ($p < 0.01$). Despite a significant increase in overall word duration for the English speakers when speaking German, L remained the same distance away from C0 as when they spoke English. This would indicate that their L tone falls much later in the word relative to segmental material when they speak English compared to when they speak German. (I will further elaborate on why this difference in alignment cannot be seen in Figure 1 later in my discussion (§4)). German speakers had similar word durations and L1 values across trial languages, so no effect of word duration was found for them in the post-hoc tests. Figure 2 shows that English speakers align their L tones later in both languages compared to the German speakers. Post-hoc tests revealed that this was only significant for the English trials ($p < 0.001$), and not for the German trials. It also shows that German speakers appear to have earlier alignment of L when speaking German compared to when speaking English, however this was not found to be significant. This contrasts the English speakers who appear to have no change in L alignment when speaking English or German, but post-hoc tests revealed a significant difference with a p-value of less than 0.001.

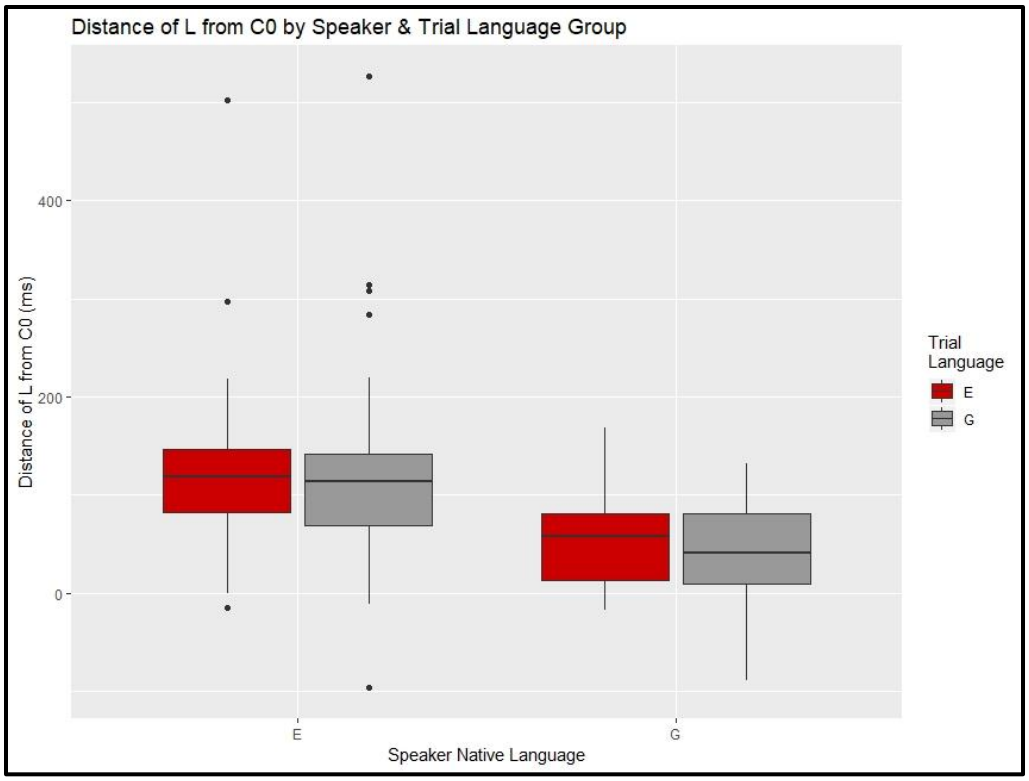


Fig. 2. Distance (in ms) of L tone from C0 (onset beginning) by speaker group, further divided by trial language. "0" on y axis represents the segmental landmark.

*3.4.2 Effect of native speaker language and trial language on L alignment to V0:* When conducting my two-way ANOVA for L2, I ran into no such issues of a word duration effect, so I was able to keep native speaker language and trial language as the only variables in my model. The interaction effect between native speaker language and trial language had a significance level of $p < 0.05$. Figure 3 shows that, like Figure 2, English speakers had later alignment in both languages compared to the German speakers. Once again, post-hoc tests revealed that this was only significant for the English trials ($p < 0.02$), and not for the German trials. The apparent later alignment of Germans when they speak German compared to when they speak English seen in Figure 3 was not found to be significant in post-hoc tests.

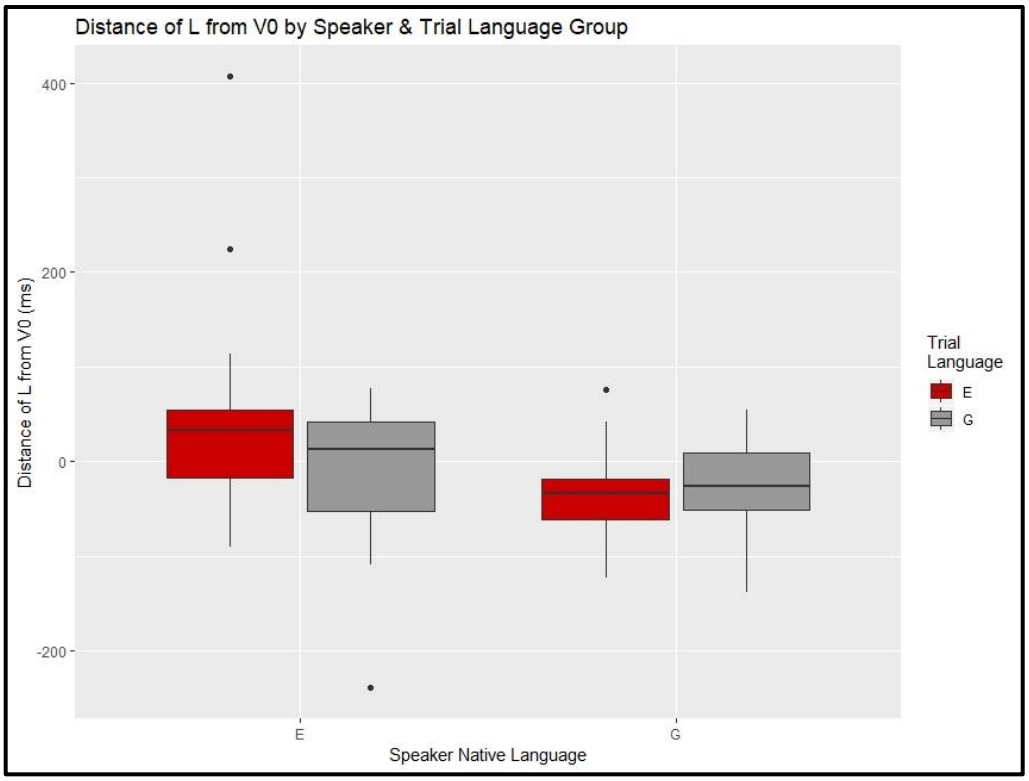


Fig. 3. Distance (in ms) of L tone from V0 (nucleus beginning) by speaker group, further divided by trial language. "0" on y axis represents the segmental landmark.

*3.4.3 Effect of native speaker language and trial language on H alignment to V1:* When conducting my two-way ANOVA for H in relation to V0, I did not find any significant effects of word duration so it was excluded as a factor from my model. I also did not find a significant main effect of trial language or any significant interaction effect between native speaker language and trial language. The main effect of native speaker language was highly significant with a p-value of less than 0.0001. Like the last two figures, Figure 4 shows that English speakers had later alignment in both languages compared to the German speakers. Post-hoc tests revealed that this was significant for the English trials ($p < 0.005$) and the German trials ($p < 0.03$). As expected, based on the ANOVA results, there were no significant differences within speaker groups across trial languages even at the post-hoc level.

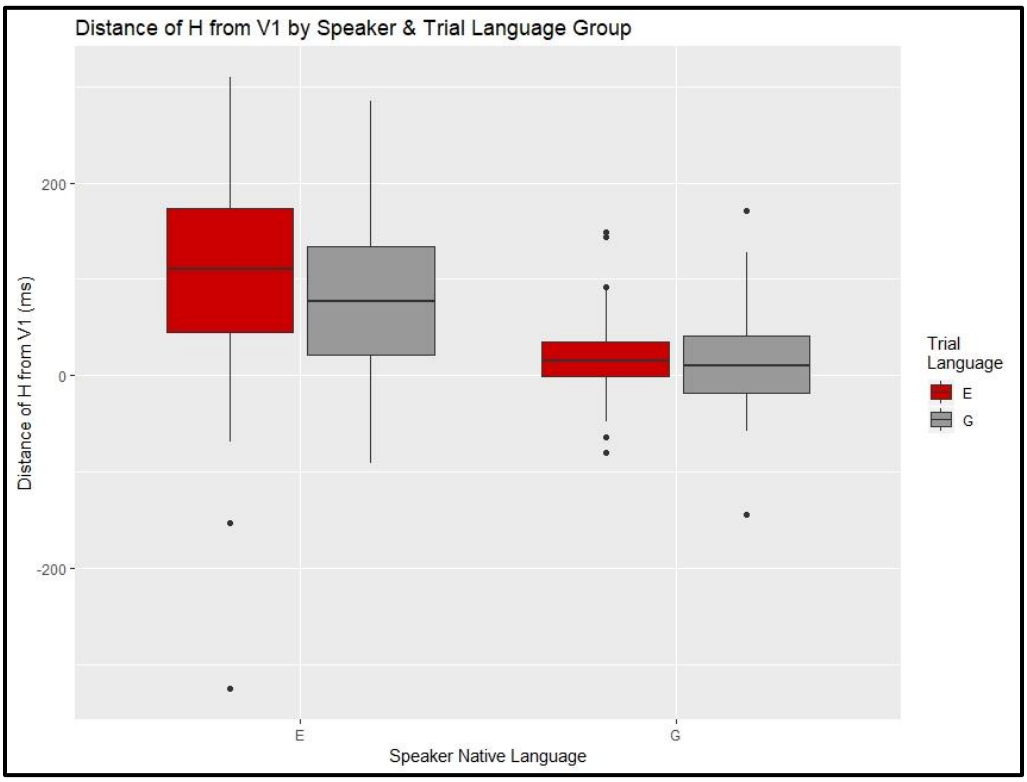


Fig. 4. Distance (in ms) of H tone from V1 (coda end) by speaker group, further divided by trial language. "0"on y axis represents the segmental landmark.

## Section 4 – Discussion

*4.1 Differences in word and segment durations*: Only total word duration was analyzed as a factor in alignment; therefore syllable, onset, nucleus, and coda duration differences will not be discussed further. The differences in total word durations generally indicate differences in speech rate across speaker groups. English speakers spoke very slowly in German, creating a large difference in duration times across their English and German trials. These are most likely due to uncertainty on the part of the Canadian English speakers when speaking German, causing a reduced speech rate. Otherwise, German speakers had a similar speech rate across both trial languages, which was comparable to the speech rate of English speakers speaking English (as evidenced by their similar total word duration times seen in Table 2). I believe that this uncertainty in the English speakers speaking German was caused by the high-level vocabulary of the German test sentences. Unfamiliarity with the words and a lack of understanding of what the sentences were saying may have caused hesitancy when they were reading.

*4.2 Tonal target $f_0$ height*: Another important aspect of alignment is whether the tonal targets had the same scaling, or pitch excursions, across speaker groups. If alignment to segmental material, and not shape or duration of contours, is what truly characterizes $f_0$ patterns (as postulated by Arvaniti et al. (1998)) then the L and H tone peaks should have similar $f_0$ values across speaker groups, or at least similar pitch excursions (difference in Hz between L and H tones). If speakers have a higher $f_0$ to reach for their H tone, they will take longer to get to that $f_0$ value, therefore the H tone may be aligned later in the word. As can be seen in Table 4, English and German speakers had very similar raw $f_0$ values for both languages, and no significant difference was found between

these values. This means that the scaling for each speaker group is similar and likely has no effect on alignment.

*4.3 Alignment of L to C0 or V0:* I can determine which landmark L is aligned to because I am making a comparison; I am stating that out of those two landmarks, it is more likely to be aligned with V0 rather than C0. As can be seen in Figure 1, L falls well within the middle of the onset for German speakers and the middle of the nucleus for English speakers. Atterer & Ladd (2004) also found that German speakers align their L tones in the middle of the onset or with the beginning of the nucleus, depending if they are speakers of Northern or Southern German dialects. Interestingly, even though my German participants were classified as Southern dialect speakers, they showed an alignment pattern closer to that of Northern German, i.e. L tone aligned within the onset. This is perhaps due to my participants originating from Middle German regions, which still have some influences from Northern German dialects, even if they are more similar overall to Southern dialects (König, 1994). My English speaker data differs greatly from that of previous research (Ladd et al., 1999; Atterer & Ladd, 2004), in that both the L and H tones are aligned far later in the syllable. Atterer & Ladd (2004) show that British English speakers align their L tones with the beginning of the onset, and their H tone lies slightly before the end of the coda. Canadian English speakers in my study align their L tones midway through the nucleus, and their H tones lie well after the end of the syllable. I believe this to be evidence for a dialectal difference in accent alignment between British and Canadian English.

I chose not to run any tests to see if H was significantly aligned to V1, because it is hard to define what “significant” alignment is without a second segmental landmark to compare it to. Is it significantly aligned if it’s within “X” ms of the landmark? How does one determine this

acceptable range? I could compare the mean alignment values of distance of L from V0 to those of H from V1. In that case, Table 3 shows that H is in fact twice as close to its landmark as L is for the German speakers, but roughly seven times as far away from the landmark as L is for the English speakers. However, if I were to run a t-test between these two sets of values, I would be looking to accept the null hypothesis. This would not be a strong signifier of significant alignment of H to V1, since it is equivalent to saying, "these two sets of values are *not different*" rather than "these two sets of values are *the same*". I note here that the H alignment for the German speakers in both languages shown in Figure 1 is inaccurate, due to the scaling of the segments at the bottom of the schematic. By looking at the grand mean distance value of H from V1 for the German speakers found in Table 3, we can see that it is positive, indicating that Germans align H after V0 rather than before (as is shown in Figure 1).

*4.4.1 Effect of native speaker language and trial language on L alignment to C0:* This difference in English speaker alignment of L in relation to C0 across trial languages cannot be seen in Figure 1 because of how the schematic was created. I relativized the L and H tones to the beginning of the syllable by subtracting the time the syllable began from the time which L and H occurred (which is equal to the distance of L or H from C0). The relative segment durations at the bottom of the graphic were the mean of the duration values from all speaker and trial language groups. The combination of this relativization and the use of overall means does not allow for the difference in English speakers' alignment for L1 across trial language to be seen. Because word duration changes across trial languages for English speakers, but distance of L from C0 does not, this would suggest an absolute alignment relative to the beginning of the word, rather than alignment with a landmark in the segmental material. However, I am doubtful that this is truly the

case because the word duration effect was much more significant than the interaction between native speaker language and trial language – p = 1e-05 for the word duration effect but p = 0.001 for the interaction effect. This difference in significance may be what causes the interaction effect to disappear when word duration is not accounted for. I also believe this word duration effect was caused by the fact that L is not aligned with C0 (as shown in §3.3), and is quite far away from that landmark, making it susceptible to changes in word duration. As for my results where, visually, I can see a difference in alignment across trial languages for the German speakers and no difference for the English speakers, but post-hoc test indicated the opposite effect: I trust Figure 2 more than the post-hoc results, since word duration is more than likely playing a role in the results of my post-hoc tests. Ultimately, the strange results produced by this word duration effect can be taken as evidence supporting alignment of L with V0 rather than C0.

*4.4.2 Effect of native speaker language and trial language on L alignment to V0:* I believe that there was no effect of word duration seen for L2 because L was shown to be aligned to V0, making it less susceptible to changes in word duration when examined in relation to this landmark. The fact that there was a significant difference between speaker groups for the English trials suggests that German speakers are transferring their German alignment patterns for L when speaking English. This is supported by Figure 3, which shows similar alignment values for German speakers in both trial languages. If they were making their patterns more like native English speakers, there would be no significant difference in post-hoc tests between speaker groups for English trials. Figure 3 would also show different alignment values across trial languages for the German speakers. The fact that no significant difference was found for English speakers across trial languages suggests only that English speakers happen to align their L tones in German close to

where native speakers align their L tone. It cannot be said that they are making changes towards native-like alignment when speaking in their L2, because there is no significant difference between their English trial and German trial L tone alignment. THerefore, I conclude that English speakers are also transferring their English alignment patterns for L when speaking German.

*4.4.3 Effect of native speaker language and trial language on H alignment to V1:* I believe that H relative to V0 was the only tonal target-segmental landmark pair to show a difference across speaker groups for both English and German trials, because the English speaker's H tones were placed so far from the H tones of the German speakers. In comparison, the L tones for each speaker group in both languages were much closer to each other. Figure 4 shows that English speakers also seem to display much greater variation, i.e. greater inconsistency, in their H tone alignment compared to the German speakers, as well as compared to their L alignment relative to either C0 or V0. This may have also contributed to making the difference between speaker groups significant. Because no significant difference was found across trial languages for both English and German speakers, I conclude that both German and English speakers transfer their alignment of H from their native language to their L2.

*4.4.4 General comments on alignment:* Interestingly, there seemed to be no effect of L2 experience on alignment, as found in Graham & Post (2017). My German speakers, who typically had 10 or more extra years of experience with their L2, transferred native alignment patterns in the same fashion as the less experienced English speakers.

## Section 5 – Conclusion

Regarding my first research question: “Do Canadian English and/or German speakers align tonal targets with segmental landmarks? If so, which ones?”, I can conclude that they do, but tonal targets tend to be consistently aligned *between* landmarks (or within segments), rather than *with* them. In their respective native languages, German speakers align their L tone within the onset and their H tone shortly after the end of the syllable, whereas Canadian English speakers align their L tone within the coda and their H tone long after the end of the syllable. Overall, Canadian English speakers were found to have a significantly later alignment of H in both English and German trials, but their later alignment of L was not significant for German trials. This suggests that English speakers’ German L tone alignment happens to be similar to native German L tone alignment. As well, Canadian English alignment patterns occur much later in the syllable compared to British English speakers, suggesting a dialectal difference.

Regarding my second research question: “Do Canadian English and/or German speakers exhibit cross-linguistic transfer of alignment patterns?”, I can conclude that they do, since for all tonal target-segmental landmark pairs there were no significant differences across trial languages within speaker groups. In other words, neither English nor German speakers had significantly different alignment patterns from their native language when speaking their L2.

## References


Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics, 26*(1), 3-25.

Arvaniti, A., Ladd, D. R., & Mennen, I. (2000). What is a starred tone? evidence from greek. *Papers in Laboratory Phonology V: Acquisition and the Lexicon,* 119-131.

Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics, 32*(2), 177-197.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook, 3,* 255-309.

Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.42, retrieved January 2018 from http://www.praat.org/

Bolinger, D. L. (1951). Intonation: Levels versus configurations. *Word, 7*(3), 199-210.

Bruce, G. (1977). Swedish accents in sentence perspective. *Working Papers in Linguistics, 12*(1)

Goldsmith, J. (1976). *Autosegmental phonology*. New York: Garland Press.

Graham, C., & Post, B. (2017). Second language acquisition of intonation: Peak alignment in American English. *Journal of Phonetics, 66,* 1-14.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent: Identifying intonational deviations by means of F0 generation and resynthesis* Inst. für Maschinelle Sprachverarbeitung der Univ. Stuttgart.

König, W. (1994). *Dtv-Atlas deutsche Sprache,* revised edition. Munich: Deutscher Taschenbuch Verlag

Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant “segmental anchoring” of f0 movements under changes in speech rate. *The Journal of the Acoustical Society of America, 106*(3), 1543-1554.

Ladd, D. R. (2008a). Fundamental concepts of the autosegmental–metrical theory. *Intonational phonology* (2nd ed., pp. 43-84). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511808814.003

Ladd, D. R. (2008b). Introduction to intonational phonology. *Intonational phonology* (2nd ed., pp. 3-42). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511808814.002

Liberman, M. Y. (1975). No title. *The Intonational System of English.,*

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32*(4), 543-563.

Pierrehumbert, J. B. (1980). No title. *The Phonology and Phonetics of English Intonation,*

Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics, 34*(3), 343-371.

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics, 33*(2), 159-197.

## Appendix A – Speech Materials

*A1.1 – German Test Sentences (test syllable in boldface)*
1. Die Ver**läng**erung der Ausleihfrist ist leider nicht möglich.
"The extension of the loan period is unfortunately not possible."
2. Die Er**nenn**ung Meiers zum Minister wurde nicht von allen Parteimitgliedern begrüßt.
"Meiers' appointment as minister was not welcomed by all party members."
3. In Er**man**gelung eines Lehrers übernahm ein Student den Unterricht.
"In the absence of a teacher a student took over the class."
4. Der **Mumm**enschanz der Kostüme rückte den Film in die Nähe eines Pop-Märchens.
"The masquerade of the costume moved the film near that of a pop-fairytale."
5. Die **Lung**entätigkeit des Patienten mußte künstlich aufrecht erhalten werden.
"The lung activity of the patient had to be artificially maintained."
6. Seine **man**gelhaften Leistungen erlaubten es ihm nicht vorzurücken.
"His defective benefits didn't allow him to advance."
7. Die **nonn**enhafte Kleidung steht ihr überhaupt nicht.
"The nun-like clothing did not suit her at all."
8. Die Ver**meng**ung der Fächer könnte zu einem Verlust führen.
"The mixing of the subjects could lead to a forfeiture."
9. Auf Ver**lang**en von Herrn Müller haben wir unser Sortiment erweitert.
"Upon request from Mr. Müller we expanded our assortment."
10. Ein **nenn**enswerter Unterschied war nicht auszumachen.
"A notable difference was hard to make out."
11. Die **moll**ige Dame bezauberte durch ihr Lächeln.
"The plump lady enchanted with her smile."
12. Die **Minn**esänger von Nürnberg waren sehr berühmt.
"The minstrels from Nürnberg were very famous."
13. Bei **Läng**engrad Null wird die Universalzeit bestimmt.
"Zero was the universal time established by longitude."
14. Die Er**lang**ung des Doktorgrades dauert immer mehrere Jahre.
"The acquisition of a PhD lasts more and more years."
15. Die be**mann**te Raumfahrt ist eine Errungenschaft des zwanzigsten Jahrhunderts.
"The manned space flight is an accomplishment of the twentieth century."

*A1.2 - German Practice Sentences*
1. Der niederträchtige Rachefeldzug ist die Handlung des dritten Aktes.
2. Nach den Warnungen des Direktors wurde Wagner vorsichtiger.

*A1.3 – German Filler Sentences*
1. Ein Nachlassen des Luftdrucks kann gefährlich sein.
2. Durch sorgfältige Planung wurde die Konferenz zum Erfolg.
3. Gegen heftige Proteste der Opposition wurde der Entwurf verabschiedet.
4. Die Ausrüstung des Bergsteigers wurde mehrfach geprüft.
5. Das Erscheinen des Hofmarschalls erstaunte die übrigen Gäste.
6. Außer dem Alter des Verdächtigen konnte nicht viel festgestellt werden.
7. Entgegen aller Erwartungen kam es am späten Nachmittag zum Eklat.

8. Das Auftreten Müllers verstörte sicherlich einige der Anwesenden.
9. Die zauberhafte Ausstattung ist ihm zu verdanken.
10. Der Elfmeterschütze des Finalspiels wurde als Held gefeiert.
11. Die Erstattung des Kaufpreises können wir nicht garantieren.
12. Die vertrauliche Anfrage erbrachte keine Ergebnisse.
13. Die außergewöhnliche Temperatur wurde in Hannover gemessen.
14. Die festgefahrene Situation sorgte für erhebliches Kopfzerbrechen.
15. Trotz angemessener Aufwandsentschädigung konnte Huber nicht zusagen.

*A1.4 – English Test Sentences (Test syllable in boldface)*
1. There was a **nom**inal fee for his services.
2. There is phe**nom**enal interest in the products.
3. She got a u**nan**imous vote for the proposal.
4. They got an a**non**ymous call from a witness.
5. He made a **lem**ony sorbet for dessert that evening.
6. She's a **min**ister's wife in the neighboring county.
7. There was an a**nom**alous reading in the data.
8. There were **mon**ogrammed sheets in the hotel rooms.
9. There is a **min**uscule chance of surviving a plane crash.
10. I need a **mon**osyllabic word for my crossword puzzle.
11. They sentenced the **mil**itant splinter group to five years.
12. You need a **min**eral and vitamin supplement to get well.
13. They charge a **min**imum rate for the use of their phone lines.
14. He took a **min**eral enriched supplement every morning.
15. They showed a **min**imal interest in what he had to say.

*A1.5 – English Filler Sentences*
1. The building in lamentable condition was torn down.
2. He was assigned a manageable workload last week.
3. The artist had malleable clay to work with.
4. The pillows made the monochromatic room much brighter.
5. There was a subliminal message in the advertisement.
6. There were many cases to be solved.
7. She found numerous insects under the rock.
8. They supplied mineralized water to the guests.
9. There were normal amounts of fluoride in the water.
10. There was monetary gain in joining the business.
11. She had rimmed glasses that were new.
12. They saw murky waters off the coast.
13. There was a marginal note near the end of the book.
14. The married couple honeymooned in Hawaii.
15. There was no running water at the cabin.

*A2.1 – Altered German Sentences*

| Item Number | Original Sentence | Altered Sentence |
|---|---|---|
| 4 | Der **Mumm**enschanz der Kostüme, das ornamentale Dekor und die synthetische Farbigkeit rückten den Film in die Nähe eines Pop-Märchens. | Der **Mumm**enschanz der Kostüme rückte den Film in die Nähe eines Pop-Märchens. |
| 8 | Die Ver**meng**ung der Fächer Medizingeschichte, Wissenschaftstheorie und Bioethik könnte zu einem globalen Niveauverlust führen. | Die Ver**meng**ung der Fächer könnte zu einem Verlust führen. |

*A2.2 – Altered English Sentences*

| | Altered English Sentences | |
|---|---|---|
| Item Number | Original Sentence | Altered Sentence |
| 6 | She's a **min**ister's wife in the Home Counties. | She's a **min**ister's wife in the neighboring county. |