9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

A Segmentation Effect in Dutch Listeners' Surface-to-Underlying Mapping of Tones

Xin Li , René Kager

Utrecht Institute of Linguistics, Utrecht University, The Netherlands
xin.li@uu.nl, R.W.J.Kager@uu.nl

Abstract

The current study investigates a) whether assimilatory tone sandhi processes, which are motivated by articulatory ease, and dissimilatory tone sandhi processes, which are not, differ in their recoverability of surface-to-underlying tone mapping for naïve Dutch listeners. It also investigates b) whether gradient and categorical underlying-to-surface sound changes in these processes lead to any difference in this mapping. Results from a series of Word Detection tasks reveal no evidence that the listeners perform surface-to-underlying tone mapping more successfully in the assimilation than in the dissimilation condition when the underlying-to-surface tone change is discrete. Also, no robust evidence is found that they perform the mapping more easily in the assimilatory condition when the underlying-to-surface alternation becomes more gradual. Consistent evidence suggests a tone segmentation effect at the surface level playing a pivotal role in the surface-to-underlying tone mapping task by the listeners. Dissimilation seems to be intrinsically related with easier segmentation whereas assimilation seems to intrinsically lead to more difficult segmentation for the Dutch listeners.

Index Terms: non-native surface-to-underlying tone mapping, assimilation vs. dissimilation, tone segmentation

1. Introduction

In connected speech, the production of a speech sound is affected by the articulatory features of its neighboring sounds [1-5]; this typically results in a different *surface* realization of the sound from its *underlying* form. The changes can either be a) assimilatory, when a neighboring sound extends its features to the coarticulated sound, e.g., a labial may change to a velar when immediately preceding another velar in Korean [6]; or b) dissimilatory, when the altered sound becomes less similar to a neighboring sound, e.g., a liquid /r/ is converted to a non-liquid /d/ before/after an adjacent liquid /l/ in Southern Bavarian German [7]. For native listeners, these surface patterns contain ambiguous information regarding the underlying form. Taking the Korean labial-to-velar assimilation rule as a hypothetical example, the surface velar-velar can underlyingly be either a labial-velar sequence, in which the underlying labial is neutralized to a surface velar, or a true velar-velar sequence. Native listeners have been amply demonstrated to be able to undo the neutralization and *map* the surface element onto the non-surface-matching underlying element, based on language-specific phonological knowledge about the underlying-to-surface (henceforth, U-to-S) process and lexical underlying representations [8-16].

Regarding naïve non-native listeners, studies have shown that they can to some extent perform surface-to-underlying (henceforth, S-to-U) mappings in assimilatory processes [17, 18], while no such evidence occurs for dissimilatory processes. By definition, naïve listeners lack knowledge of phonological processes and lexical representations. The question arises how they accomplish mappings in assimilatory processes, and what are the representations onto which surface forms are mapped.

The different status of nonnative mappings in assimilatory and dissimilatory processes could reside in their different motivations. Assimilatory processes are mostly believed to be motivated by "articulatory ease" [19, 20]. Dissimilatory sound changes were proposed by Ohala [21, 22] to occur as listeners are under pressure to preserve the underlying segment in the surface form, and misattribute a feature that is intrinsic to a surface form to coarticulation, causing a "hyper-correction". Motor Theory [23, 24] and Direct Realist Theory [25-28] agree that acoustic events are perceived as being caused by articulatory gestures. Based on this idea, we hypothesize that what non-native listeners might be doing in S-to-U mappings is to map a surface form onto a hypothetical underlying sound, invoking articulatory knowledge that is independent of native language experience, interpreting non-native sounds in terms of articulator settings. This would predict that S-to-U mapping for assimilations uses language non-specific knowledge based on articulatory ease, which is accessible for naïve listeners; while mappings for dissimilations should be language-specific, as such processes crucially refer to lexical representations that are unavailable for naïve listeners.

Extending the mapping hypothesis for assimilatory process based on articulatory ease, it could be that non-native listeners can recover the underlying representation only in case this is to some extent acoustically recoverable from the surface. The change from underlying to surface forms can be *categorical*, when the articulatory characteristic of the underlying element is *completely lost* in the production of the surface form, e.g., in Korean labial-to-velar assimilation, surface labials neutralize with underlying velars with converging F2 and F3 values [29]. Alternatively, the change can be *gradient* when an articulatory gesture of an underlying element is *partially lost* at the surface, e.g., assimilated coronals in English coronal place assimilation have formant values intermediate between those of coronals and non-coronals [30]. Hence, we may assume that gradient assimilations, leaving phonetic traces of the underlying sound at surface, will allow naïve listeners to do the S-to-U mapping.

Empirical evidence supports the hypothesis that naïve non-native listeners can perform S-to-U mapping for assimilations, with mixed evidence on whether such mappings crucially rely on gradience. Most studies found that naïve listeners were able to recover underlying sounds successfully only for gradient

 10.21437/SpeechProsody.2018-29

assimilations. For instance, [17] observed that naïve English listeners performing a mapping task for a gradient Hungarian voicing assimilation process were facilitated in monitoring a target underlying sound in its surface form in a viable context, as compared to an unviable context; whereas listeners did not show facilitating effects for Korean categorical labial-to-velar assimilation. An ERP study [18] tested Dutch listeners on Hungarian /l/-to-/r/ assimilation, and found listeners accepted the assimilated segment as its underlying counterpart more in a viable than unviable context when they heard a partially assimilated /r/. One study suggests that naïve listeners relate surface and underlying sounds in categorical assimilations. [11] tested naïve French listeners on English place assimilation as well as naïve English listeners on French voicing assimilation, and found that both groups detected the underlying word more successfully in viable than in unviable contexts, using spliced neutralized phonemic contrasts.

Regarding dissimilatory processes, no previous study has investigated whether naïve non-native listeners can perform S-to-U mappings, as far as we know. Whether S-to-U mapping in dissimilatory processes requires native language experience inaccessible to non-native listeners, as Ohala [21, 22] implies, remains unclear. The current study aims to investigate a) whether assimilatory and dissimilatory processes differ in their recoverability in S-to-U mapping for naïve nonnative listeners. It also investigates b) whether gradient and categorical sound changes lead to differences in mappings. We use lexical tone and tone sandhi processes to investigate these issues. By using lexical tone, we are able to adopt a group of non-tone listeners that is maximally remote from native listeners, which offers the strongest test for the language-general mapping hypothesis for assimilatory processes. An artificial tone language will be constructed, manipulating two dimensions of U-to-S changes: a) assimilatory versus dissimilatory and b) gradient versus categorical. Dutch listeners will be used as the naïve non-tone language listeners.

2. Experiments

The Word Detection paradigm was successfully used by [11] to test native and naïve listeners on their S-to-U mapping in assimilatory processes. Here we apply this paradigm to tone sandhi processes, to test whether naïve Dutch listeners can map a surface sandhied tone onto a target (underlying) tone. Target words (in their underlying tonal shape) were presented auditorily and followed by test items containing a target word (now appearing in a gradient vs. categorical sandhied tone shape, and embedded in an assimilatory/dissimilatory context). Participants had to decide whether they spotted the target word in the test items. A comparison of detection rates of target words in a) assimilatory vs. dissimilatory contexts and in test words containing b) gradient vs. categorical sandhied tones will demonstrate how Dutch listeners make use of the two factors in a S-to-U tone mapping task.

2.1. Experiment 1

Experiment 1 included categorical U-to-S changes, either assimilatory or dissimilatory. Based on the hypothesized language-unspecific S-to-U mapping in assimilatory processes and language-specific mapping in dissimilatory processes, our prediction was that naïve Dutch listeners would perform more successfully S-to-U mapping in an assimilatory process than in a dissimilatory process.

2.1.1. Stimuli

A pair of tone sandhi rules involving categorical changes and differing in assimilatory/dissimilatory nature were created in a made-up language. Sandhi 1: 44.24 $\rightarrow$ 42.24[1]; Sandhi 2: 44.42 $\rightarrow$ 42.42. In Sandhi 1 the offset of a high-level 44 in the first syllable is lowered to the same pitch as the onset of the following 24, resulting in a categorical assimilation process; Sandhi 2 is a categorical dissimilation process, as the offset of a high-level tone in the first syllable deviates far from the subsequent onset of a high-falling 42, instead of approaching it. The sandhi rules both involved a change from an underlyingly high-level tone 44 to a surface high-falling tone 42.

Four target words carrying the underlying tone 44 were selected: /ba44/, /bi44/, /de44/, and /go44/. The target words were spoken by a 27-year-old female Mandarin speaker A. Tokens were digitized at 16kHz and 16 bits, and normalized in duration (450 ms) and intensity. Four matched test words with tone 42 ([ba42], [bi42], [de42], [go42]) and two context words [du24] and [du42] were recorded by a 28-year-old female Mandarin speaker B. Test words and context words were first normalized; then each of the 4 test words was concatenated with the 2 context words with 70 ms in between, creating 8 disyllabic test items (e.g. [ba42du24]/[ba42du42]). Test items were repeated 8 times and randomly assigned into 2 blocks.

In order to draw participants' attention to the tonal shape of words, 8 additional test items were constructed as practice trials. The practice trials contained a tonally changed test item (e.g. [ba24du22]) and a tonally unchanged test item (e.g. [ba44du24]) for each of the 4 target words. None of the practice items carried the same tone pattern with the 2 experimental conditions. All recordings and editing were done using Audacity and Praat [32].

2.1.2. Participants

Thirty Dutch participants with no previous exposure to a tone language were recruited. All were aged between 18 and 30, and all self-reported normal hearing. Every participant did the two conditions (assimilatory 42.24/dissimilatory 42.42) in one experiment. The order of trials within a block and the order of the blocks were randomized for each participant.

2.1.3. Procedure

Experimental trials were pairs beginning with the presentation of the monosyllabic target word (e.g. [ba44]), followed after 1000 ms of silence by a disyllabic test item (e.g. [ba42du24]). The intertrial interval was 4000 ms. Participants were requested to press a "yes" button when they thought the target appeared in the test item with the same melodic shape, and press "no" if otherwise. In each trial they were allowed 2000 ms after the offset of the target word to make their response. Missed trials were repeated at the end of the trials.

In the practice session, practice trials were played to the participants in a loop. Visual "correct/incorrect" feedback was provided after each response. The practice session would automatically end when a listener performed 7 trials in a row

[1] The digits indicate the rough heights of pitch using [31]'s 5 level representation, with 1 representing the lowest pitch level in the listener's pitch range and 5 the highest; the dot is used to separate tones.

with ≤ 2 errors. Then they proceeded to the test session, during which responses were collected without feedback. The experiment protocol was run via ZEP [33]. Detection (yes/no) for each test item was used as the main measure.

2.1.4. *Results and discussion*

A generalized linear mixed-effects model was constructed to analyze the detection data. PARTICIPANT and ITEM were included as random effects. Fixed effect of SURFACE FORM (assimilatory 42.24/dissimilatory 42.42) was added to the model and proved to significantly improve the fit of the model (p= .001 **). The model revealed significantly lower detection of the underlying tone 44 in the surface form 42.24 than in the surface form 42.42 (p= .001 **).

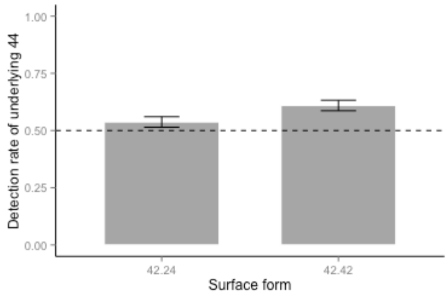

Figure 1: *Detection rate of the underlying 44 in the surface forms 42.24 and 42.42 in the Dutch participants.*

It turned out that the Dutch listeners demonstrated a word detection pattern opposite to our hypothesis of better mapping in the assimilatory condition. Instead, the listeners detected the underlying tone 44 more frequently in the 42.42 surface form, involving a categorical dissimilatory change from underlying form, than in the 42.24 surface form, which included discrete assimilatory change. The cause of the observed unexpected better mapping in the dissimilatory condition will be discussed together with the results of Experiment 2 in the next section.

The finding that the S-to-U mapping did not benefit from the assimilatory condition in Experiment 1 might be due to the categorical change from the underlying tone 44 to the surface tone 42, which left no acoustic traces of the target (underlying) tone in the surface form for the Dutch listeners to perform the mapping. Experiment 2 aims to investigate the same mapping when the change becomes more gradual.

2.2. Experiment 2

Experiment 2 changed the surface tone to 43, i.e. a less steeply falling tone that is intermediate between 42 and 44. We hypothesized that when the change is more gradient, better S-to-U mapping might be observed in the assimilatory condition.

2.2.1. *Stimuli*

The same stimuli as in Experiment 1 were used, except that tones in monosyllabic test words were changed to 43, and accordingly, tone sequences on the disyllabic test items were altered to 43.24 and 43.42.

2.2.2. *Participants*

Another thirty Dutch participants were recruited. They met the same criteria as in Experiment 1.

2.2.3. *Procedure*

The procedures were the same as in Experiment 1.

2.2.4. *Results and discussion*

A generalized linear mixed-effects model revealed no significant difference between the detection rates of the 43.24 and the 43.42 conditions. (p> .5).

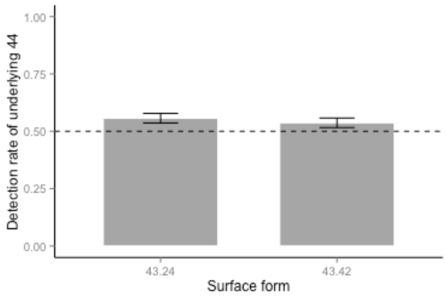

Figure 2: *Detection rate of the underlying 44 in the surface forms 43.24 and 43.42 in the Dutch participants.*

The results showed that the Dutch listeners did not succeed in significantly better S-to-U tone mapping in the assimilatory condition, even when the U-to-S tonal change was more gradual. A noticeable trend of better mapping in the assimilatory condition than in the dissimilatory condition was observed, but the difference was statistically insignificant.

In the S-to-U tone mapping tasks in Experiment 1 and 2, the Dutch listeners seemed to be mainly confronted with a segmentation task at the surface tone level.

In Experiment 1, the Dutch listeners seemed to be facilitated in segmenting the surface tone 42 from its following tonal context in the dissimilatory 42.42 condition; whereas they seemed to have experienced increased difficulty in segmenting the surface tone 42 from its immediately following tonal context 24 in the assimilatory 42.24 condition.

The role of dissimilation as a potential segmentation marker was already noticed by [20], who observed that cross-linguistically, dissimilation phenomena occur mostly across concatenated morphemes, and suggested that dissimilation processes apply to “preserve the distinctiveness of the stem-affix relationship” and to “make stem boundaries more prominent”. In our Experiment 1, the Dutch listeners may have also set up a hypothetical morpheme boundary in the dissimilatory 42.42 condition, because of the F0 disagreement between the surface tone and its tonal context, and thus may have segmented the surface tone 42 easier from its context; whereas the f0 agreement between the surface tone 42 and the contextual tone 24 in 42.24 may have lead to more difficult segmentation of the surface tone 42. Evidence from previous studies suggests that an f0 jump can be used to suggest boundaries. For example, [34] found a clear effect of pitch reset on the detectability of discourse boundaries and on subjects’ ranking of boundary hierarchy; [35] found that silent intervals bounded by tokens of widely differing pitch are heard as longer than those bounded by tokens closer in pitch.

A different type of segmentation cue could have come from double occurrence (repetition) of tone. In the 42.42 surface condition of Experiment 1, the repetition of tone 42 might have contributed to the easier segmentation of the surface tone 42 from its context. Studies such as [36] noted the underrepresentation of syllable repetition in lexicon in many languages. [37] observed that the co-occurrence of pairs of identical consonants was used as a segmentation cue for Dutch listeners. The restriction on repetition of identical tones is wide spread in the lexicon and surface phonology of many languages, e.g. OCP-high tone in Bantu languages [38].

In Experiment 2, the contrast between f0 agreement and disagreement was eliminated in the surface forms 43.24 and 43.42; meanwhile, repetition was avoided in the surface tone and its tonal context. Hence, it may be assumed that the segmentation of the surface tone sequence in the two conditions was equally clear/unclear to the Dutch participants. As segmentation played no role in Experiment 2, the hypothesized better mapping in the assimilatory (compared to dissimilatory) condition should have re-emerged. We indeed observed this tendency as a trend in the results, although this effect was not statistically significant.

The current results left it unclear whether the facilitated segmentation for the Dutch listeners in the 42.42 surface form in Experiment 1 was due to a) f0 disagreement on the tone boundary or b) repetition of the 42 tone. Experiment 3 was included to disentangle these two possible effects.

2.3. Experiment 3

Based on Experiment 1, Experiment 3 added a third tone condition 42.44322 with f0 disagreement intact (+ f0 disagreement) while repetition removed (- repetition), to allow its comparisons with the 42.24 condition (- f0 disagreement, - repetition) and with the 42.42 condition (+ f0 disagreement, + repetition). The predictions are: a) if the 42.44322 surface elicits the same detection rate as 42.24, both significantly lower than 42.42, then it should be repetition that makes the segmentation easier for the listeners; b) if 43.44322 produces the same detection rate as 42.42, both significantly higher than 42.24, that f0 disagreement should be the main cause of easier segmentation; c) if the detection of 44.44322 is intermediate between 42.24 and 42.42, then both f0 disagreement and repetition are contributing to easier segmentation.

2.3.1. Stimuli

The stimuli in Experiment 1 were included in the current task, and a new test item condition 42.44322 was added. The tone shape 44322 was created by twisting the 42 tone in Praat [32] while keeping the duration intact.

2.3.2. Participants

Another thirty Dutch participants were recruited. They met the same criteria as in Experiment 1& 2.

2.3.3. Procedure

The procedure was the same as in Experiment 1 and 2.

2.3.4. Results and discussion

A generalized linear mixed-effects model revealed that SURFACE FORM (42.24/42.42/44.322) was a significant predictor of the model (p= .027 *). A post-hoc paired comparison between conditions revealed: a) significantly higher detection rate in the 42.42 condition than the 42.24 condition (p= .03 *); b) slightly but not significantly higher detection rate in the 44.322 condition than in the 42.24 condition (p= .10); c) no difference between the detection rates of the 43.24 and the 43.42 conditions. (p= .54).

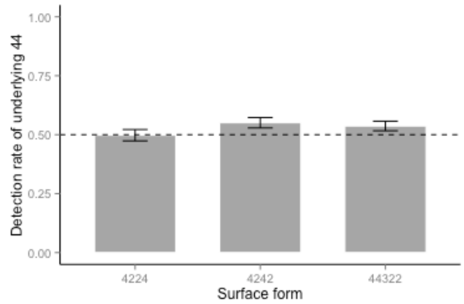

Figure 3: *Detection rate of the underlying 44 in the surface forms 42.24, 42.42 and 44322 in the Dutch participants.*

The significantly higher detection rate in the surface form condition 42.42 than in the surface form condition 42.24 was replicated in the current experiment, which further confirmed that the segmentation of the surface tone 42 was facilitated for the Dutch listeners in the 42.42 tone sequence.

The surface form 42.44322 elicited a detection rate comparable with the condition 42.42. This persuaded us to believe that the repetition of the tone 42 was unlikely to be the main cause of easier tone segmentation. The 42.44322 condition elicited slightly but not significantly higher detection rate than the 42.24 condition. It did not confirm us that the f0 jump necessarily lead to easier segmentation for the Dutch listeners, though it was more likely to be the major cause.

3. General discussion and conclusions

No evidence was found suggesting that the Dutch participants succeeded in S-to-U tone mapping more successfully in the assimilation condition than in the dissimilation condition, when tonal changes were discrete. When the changes became gradual, still no robust evidence was found supporting that the S-to-U mapping is easier in the assimilatory condition, although a trend of the hypothesized difference is noticed.

Instead, cumulative evidence was found suggesting that tone segmentation may play a cardinal role in Dutch listeners' surface-to-underlying tone mapping task, and may prevent access to underlying representations. In particular, the listeners seemed to have segmented the surface tone from the disyllabic sequence more easily in the dissimilation condition 42.42; whereas they experienced increased difficulty in segmenting the surface tone from its adjacently following tonal context in the assimilation condition 42.24. Whether it is f0 jump in the tone sequence or tone repetition that caused segmentation ease for Dutch participants remains unclear; however, the f0 jump is more likely to be the main cause. Dissimilation by its nature implies discontinuity, and hence may have an intrinsic relationship with easier segmentation, as opposed to assimilation, which implies continuity and may intrinsically lead to more difficult segmentation.

The task we employed in the current study asks listeners to first attend to the target in isolation and then retrieve the target in the context, which invites attention to segmentation. Future studies exploring the S-to-U mapping problem by naïve listeners may consider this segmentation effect and use experiment methods that minimally allow it to play a role.

4. References

[1] Beddor, P.S. and R.A. Krakow, Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *The Journal of the Acoustical Society of America*, 1999. 106(5): p. 2868-2887.

[2] Fowler, C.A., Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 1980. 8(1): p. 113-133.

[3] Hammarberg, R., The metaphysics of coarticulation. *Journal of Phonetics*, 1976. 4: p. 353-363.

[4] Lahiri, A. and H. Reetz, Underspecified recognition. *Laboratory phonology*, 2002. 7: p. 637-675.

[5] Xu, Y., Production and perception of coarticulated tones. *The Journal of the Acoustical Society of America*, 1994. 95(4): p. 2240-2253.

[6] Kim-Renaud, Y.-K., Korean consonantal phonology. 1974, *Ph.D. dissertation*, University of Hawaii.

[7] Hall, T.A., Liquid Dissimilation in Bavarian German. *Journal of Germanic Linguistics*, 2009. 21(1): p. 1-36.

[8] Chen, A. and R. Kager. The perception of lexical tones and tone Sandhi in L2: success or failure. in *Congress of Phonetic Sciences XVII proc*. 2011.

[9] Chen, A., L. Liu, and R. Kager, Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences*, 2015. 48: p. 62-69.

[10] Coenen, E., P. Zwitserlood, and J. Bölte, Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, 2001. 16(5-6): p. 535-564.

[11] Darcy, I., et al., Phonological knowledge in compensation for native and non-native assimilation. *Variation and gradience in phonetics and phonology*, 2009. 14.

[12] Gaskell, M.G. and W.D. Marslen-Wilson, Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human perception and performance*, 1996. 22(1): p. 144.

[13] Gaskell, M.G. and W.D. Marslen-Wilson, Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1998. 24(2): p. 380.

[14] Gaskell, M.G. and N.D. Snoeren, The impact of strong assimilation on the perception of connected speech. *Journal of Experimental Psychology: Human Perception and Performance*, 2008. 34(6): p. 1632.

[15] Li, X., R. Kager, and W. Gu. Surface vs. Underlying Listening Strategies for Cross-Language Listeners in the Perception of Sandhied Tones in the Nanjing Dialect. in *The 5th International Symposium on Tone Aspects of Languages*. 2016.

[16] Mitterer, H., S. Kim, and T. Cho, Compensation for complete assimilation in speech perception: The case of Korean labial-to-velar assimilation. *Journal of Memory and Language*, 2013. 69(1): p. 59-83.

[17] Gow, D.W. and A.M. Im, A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, 2004. 51(2): p. 279-296.

[18] Mitterer, H., et al., The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science*, 2006. 30(3): p. 451-479.

[19] Grammont, M., Traité de phonétique. 1950: Librairie Delagrave.

[20] Johnson, L., Dissimilation as a Natural Process in Phonology. 1973.

[21] Ohala, J., J. 1981. The listener as a source of sound change. *Papers from the Parasession on Language and Behavior*: p. 178-203.

[22] Ohala, J.J., The phonetics of sound change. *Historical linguistics: Problems and perspectives*, 1993: p. 237-278.

[23] Liberman, A.M., et al., Perception of the speech code. *Psychological review*, 1967. 74(6): p. 431.

[24] Liberman, A.M. and I.G. Mattingly, The motor theory of speech perception revised. *Cognition*, 1985. 21(1): p. 1-36.

[25] Fowler, C.A., An event approach to the study of speech perception from a direct-realist perspective. *Status Report on Speech Research*, edited by IG Mattingly and N. O'Brien, Haskins Laboratories, New Haven, CT, 1986: p. 139-169.

[26] Fowler, C.A., Listeners do hear sounds, not tongues. *The Journal of the Acoustical Society of America*, 1996. 99(3): p. 1730-1741.

[27] Fowler, C.A., Compensation for coarticulation reflects gesture perception, not spectral contrast. *Perception & Psychophysics*, 2006. 68(2): p. 161-177.

[28] Fowler, C.A. and J.M. Brown, Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Attention, Perception, & Psychophysics*, 2000. 62(1): p. 21-32.

[29] Jun, J., Place assimilation. *Phonetically based phonology*, 2004: p. 58-86.

[30] Gow, D.W., Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, 2003. 65(4): p. 575-590.

[31] Chao, Y., A system of tone letters. *Le Maitre Phonetique*, 1930.

[32] Boersma, P., Weenink, D. , Praat: Doing phonetics by computer. 2010.

[33] Veenker, T.J.G., The Zep Experiment Control Application (Version 1.14.5) [Computer software]. Beexy Behavioral Experiment Software. Available from http://www.beexy.org/zep. 2017.

[34] Lin, H.-Y. and J. Fon. The role of pitch reset in perception at discourse boundaries. in *17th International Congress of Phonetic Sciences*. Hong Kong. 2011.

[35] Brugos, A. and J. Barnes. Effects of dynamic pitch and relative scaling on the perception of duration and prosodic grouping in American English. in *Speech Prosody*. 2014.

[36] Monaghan, P. and W. Zuidema. General purpose cognitive processing constraints and phonotactic properties of the vocabulary. in *Workshop on the Evolution and Phonetic Capabilities: Causes, Constraints and Consequences*. Retrieved from http://www.lancaster.ac.uk/staff/monaghan/papers/monaghan_zuidema_15_icphs.pdf. 2015.

[37] Boll-Avetisyan, N. and R. Kager, OCP-PLACE in speech segmentation. *Language and Speech*, 2014. 57(3): p. 394-421.

[38] Meyers, S., OCP effects in Optimality Theory. *Natural Language & Linguistic Theory*, 1997. 15(4): p. 847-892.