

Novel Popout in English Phoneme Monitoring

Jeff Berry

May 3, 2007

1. Introduction

This paper presents findings from a phoneme monitoring experiment conducted on native English speakers. Phoneme monitoring consists of presenting a stimulus to a subject and asking the subject whether the stimulus contains a certain phoneme or allophone. The subject's response time is recorded, and can be used to argue for different theories of speech perception and processing. As this paper will show, response times are sensitive to a variety of factors. This paper will focus specifically on the role phonological assimilation rules play in phoneme monitoring tasks.

Two types of assimilation rules will be discussed: regressive and progressive. Regressive assimilation rules are fairly common in the world's languages. The rule is called regressive because a later segment affects an earlier one. An example from Japanese will illustrate this phenomenon, in the form of nasal place assimilation. Nasal place assimilation is very common cross-linguistically, making it a particularly interesting case. In Japanese, the nasal in *san* 'three' undergoes regressive place assimilation, taking on the place of articulation of the following consonant:

- (1) *Japanese regressive nasal place assimilation*
- | | | | |
|----------------|---|------------|----------|
| (a) /sangatsu/ | → | [sangatsu] | 'March' |
| (b) /sanban/ | → | [samban] | 'third' |
| (c) /sanju/ | → | [sanju] | 'thirty' |

Besides regressive assimilation, many languages exhibit progressive assimilation, where an earlier segment affects a later one. A clear example of progressive assimilation is German fricative assimilation. As (2) shows, the palatal fricative [ç] becomes the velar fricative [x] when it follows a back vowel.

- (2) *German progressive fricative assimilation*
- | | | | |
|------------|---|--------|----------|
| (a) /lɪçt/ | → | [lɪxt] | 'light' |
| (b) /laçt/ | → | [laxt] | 'laughs' |

This paper will discuss how violations of these assimilation rules affect response times in phoneme monitoring tasks. Specifically, it will show a novel pop-out effect when English progressive rules are violated, similar to the effect found in German by Weber 2001. The effect is referred to as novel popout because novel sequences have faster reaction times than non-novel sequences. A novel sequence is defined here as a combination of sounds that never occurs in the language, and is thus novel, for example [ɣ] in German.

The remainder of the paper is organized as follows: section 1.1 will discuss previous studies that have examined assimilation rules involving phoneme monitoring. Section 2 will introduce the current experiment. Section 3 will describe the method used in this experiment, section 4 will give the results, and section 5 will be the discussion and conclusion.

1.1. Previous Studies

1.1.1. Gaskell & Marslen-Wilson 1996, 1998

Of the two studies conducted by Gaskell and Marslen-Wilson, the first (1996) deals with the effects of auditory assimilation violations on visual lexical recognition, while the second (1998) involves phoneme monitoring. The first experiment used cross-modal priming. The prime consisted of a sentence produced auditorily, in which the last word was truncated, leaving the prime word at the end of the sentence, which had its final consonant assimilated to the first consonant of the removed word in some instances. For example, the word *wicked* was originally followed by *prank*, causing the [d] to assimilate to [b], producing the prime word [wɪkɪb]. In other instances, the assimilation did not take

place, producing a second prime word [wɪkɪd]. The prime was immediately followed by the target word *wicked*, which was presented visually on the computer screen. Subjects had to make a lexical decision on the visually presented word. Both the assimilated and unassimilated words produced a strong priming effect. Gaskell and Marslen-Wilson state (1998:151), “Experiment 1 showed a strong priming effect for both changed and unchanged prime words, with little evidence of a mismatch effect for the distorted words.”

In Experiment 2, the sentences were not truncated, but the final words were changed, to cause a mismatch between the assimilated prime word and the following word. For example, [wɪkɪb] was followed by *game* rather than *prank*, which presented a violation of the assimilation rule. As in Experiment 1, the target word *wicked* was displayed visually immediately following the prime, and subjects were asked to make a lexical decision. In this experiment, there was a difference in response times. When the assimilation was misapplied in the priming sentence, subjects were slower to make a lexical decision on the target. From this study, it appears that lexical decision is only affected by assimilation if the rule is violated, as in the case of *[wɪkɪb geɪm].

In a follow-up study, Gaskell and Marslen-Wilson (1998) use phoneme monitoring to test the effects of assimilation rules on processing. The stimuli were similar to those in the 1996 study, in that they involved optional assimilation of a final consonant of one word to the first consonant of the second word. For example, subjects were presented with the following stimuli: [fʌn kæmp], [fʌŋ kæmp], and [fʌm kæmp]. Subjects were asked to monitor for [k]. Gaskell and Marslen-Wilson found that there was no difference in response time for [fʌn kæmp] and [fʌŋ kæmp] (mean RTs of 476ms and

480ms respectively), where there has been either no assimilation, or the assimilation is phonologically viable. However, response times to [fʌm kæmp] were slower (mean 536ms) when the assimilation rule was violated, producing an unviable sequence. To summarize, lawful application of an assimilation rule did not effect response times, but misapplication significantly slowed response times.

1.1.2. Otake, Yoneyama, Cutler & van der Lugt 1996

In their study on Japanese moraic nasals, Otake et al. (1996) conducted phoneme-monitoring experiments involving violations of obligatory nasal place assimilation rules. For stimuli, they took words such as *tombo* ‘dragonfly’, and *kondo* ‘this time’ that display the obligatory nasal place assimilation, and spliced in different nasals. When these stimuli were presented to Japanese listeners, response times in monitoring for the post-nasal consonant were significantly slower when the nasal did not match the place of articulation. They state (1996:3839), “The mean RT to consonants following nasals in their original contexts was 544 ms, to consonants following cross-spliced nasals 592 ms, and this difference was significant ($F_1[1,39]=25.95$, $p<0.001$; $F_2[1,52]=7.11$, $p<0.02$).” From this study and the Gaskell and Marslen-Wilson studies discussed above, it is clear that when regressive assimilation rules are violated, they slow down response times.

1.1.3. Weber 2001

Weber 2001 argues that violations of regressive assimilation rules produce inhibitory effects in phoneme monitoring because they violate strong expectations for the following sound. For example, when a German speaker hears an [ŋ], the phonotactics of

German lead the person to expect a following sound from a very limited set. Weber states (2001:40), “In /fɛŋ-/, for instance, /ŋ/ could have been followed by only three consonants, /k/, /s/, and /t/. Thus, the regressive assimilatory effect strongly restricted what the following segment might be.” The opposite holds for progressive assimilations, where expectations for following sounds are very low. To take another example from German, a front vowel can be followed by many sounds, including [ç]. However, a front vowel does rule out a subsequent [x]. Weber argues that there is thus a difference between regressive and progressive assimilations, in that regressive rules produce strong expectations, whereas progressive rules produce only weak expectations. This leads to the natural question of how progressive assimilations affect processing.

To explore this question, Weber’s experiment 1 presented Dutch stimuli to German speakers. Dutch does not have the progressive assimilation rule outlined in (2) above, and can therefore have a front vowel followed by the velar fricative [x]. German speaking participants were instructed to monitor for [x] while listening to the Dutch stimuli. Surprisingly, response times were faster for items that violated the German progressive assimilation rule as Table 1 shows (from Weber’s Table 2-1):

Table 1. Weber’s (2001) results for Experiment 1. Surprisingly, the items that violated the German progressive rule were responded to faster.

Measure	Back vowel [bɑxt]	Front vowel *[bɛxt]
RTs	498	470
Errors	0.8%	1.1%

When these same stimuli were given to Dutch speakers, no effect was found.

Weber’s Experiment 5 revealed another interesting factor: novelty of the resulting sequence of sounds. A novel sequence is a group of two sounds that do not occur in the language. An example of a novel sequence in German is [ɛx], which is responded to

faster than words that do not have a violation, as in Experiment 1. However, experiment 5 showed that if a violation of a progressive assimilation rule produces a non-novel sequence, response times are *slower* than words without a violation. This experiment also tested German fricative assimilation, but this time the violation produced a back vowel followed by [ç], which Weber argues is a non-novel sequence since it occurs across word boundaries as in *Rosa Chinese*. Table 2 shows that the sequence *[aç] is slower than non-violating words (from Weber’s Table 2-5):

Table 2. Weber’s (2001) results for experiment 5. Contrary to experiment 1, the violation was responded to slower than the non-violating item.

Measure	Front vowel [bi:çel]	Back vowel *[ba:çəl]
RTs	547	564
Errors	1.7%	0.8%

From this, Weber attributes the faster response times in experiment 1 to a phenomenon known as “novel popout”. The term novel popout comes from work done on visual perception that shows that novel items are processed faster than ordinary ones. Johnston and Schwarting (1996:211) argue that

“because organisms already know their familiar habitats, it would be a waste of their time and energy to engage in detailed physical analyses of these habitats every time they are encountered...expected inputs generate a high degree of conceptual processing but a low degree of physical processing. A by-product of the suppressed physical processing of expected inputs...is the enhanced physical processing of any unexpected inputs in their midst (i.e., novel popout).”

Thus the novel sequence [ɛx] produces a faster response time than the non-novel sequence [aç].

To summarize, Weber (2001) argues that facilitation occurs only when there is a violation of a progressive rule that produces a novel sequence. A novel sequence produced by a regressive rule violation does not speed response times, as can be seen the results of the Otake et al. (1996) study, where novel sequences such as [nb] had slower response times.

2. Current Experiment

The current experiment was designed to test for novel popout effects in English, similar to those found in Weber's experiment 1. In order to test Weber's predictions, a progressive assimilation rule was needed that would produce a novel sequence. English has very few progressive assimilation rules—the author could only think of the progressive voicing assimilation rule for the plural and past tense morphemes. Although violations of these rules do not produce novel sequences ([tz] occurs across word boundaries: *the cat zipped up the stairs*), the case of the plural morpheme was chosen. The difference in response times to the non-violating sequence [ts] was compared to the violating sequence [tz]. In order to test whether novel popout could occur separately from the context of an assimilation violation, a novel sequence was also tested: [tɿ] was compared to [ɿt].

3. Method

3.1. Participants

22 undergraduates from the INDV 101 course participated for extra credit. 4 of the subjects' results were excluded because they were not native English speakers, making a total of 18 native English-speaking participants.

3.2. Stimuli

For this experiment, there were 165 non-words, divided into three groups of 55 non-words each. All of the stimuli were constructed using Mono-nonsense, an extensive list of monosyllabic nonsense words that obey the phonotactics of English, made by Mike Hammond. Each of the three groups were composed of the following:

- A. 10 items with the target sound following [t],
- B. 10 items with the target sound in other phonological environments,
- C. 30 filler items that did not contain the target sound,
- D. 5 practice items, some of which contained the target sound.

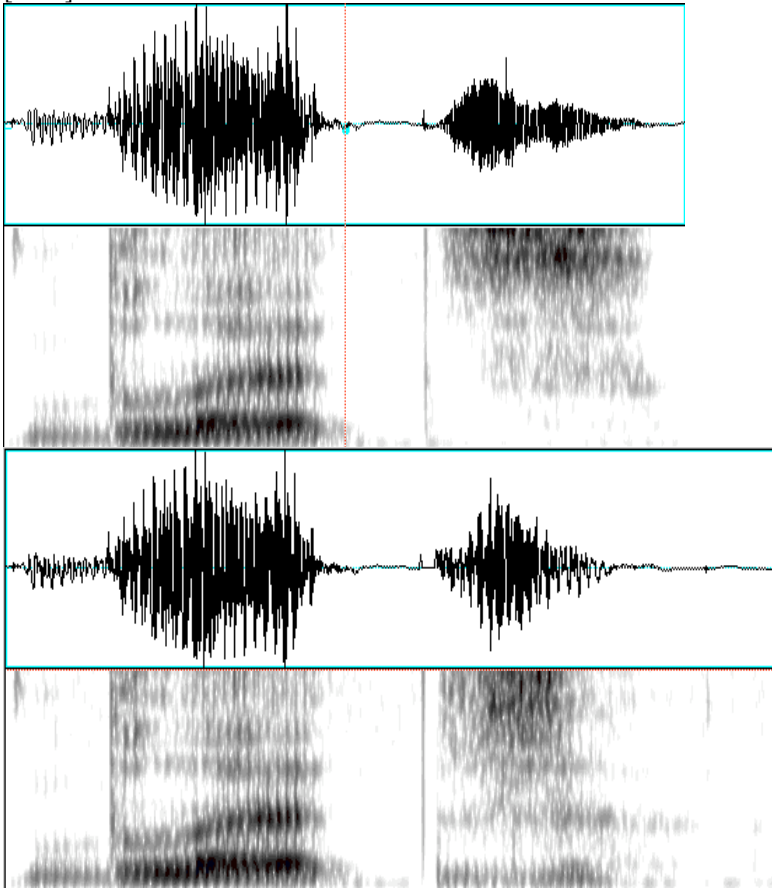
There were three target sounds (hence the three groups): [s], [z], and [ŋ]. Table 3 shows the A and B items for each group.

Table 3. List of A and B items for each group. The A items (left column) have the target sound following [t], the B items (right column) have the target sound in other phonological environments.

[s] group		[z] group		[ŋ] group	
[ts]	[s]	[tz]	[z]	[tŋ]	[ŋt]
[blɛts]	[bʊst]	[blɛtz]	[glɪnz]	[blɛtŋə]	[blɛŋtə]
[tʃʌts]	[dæps]	[tʃʌtz]	[hɪvz]	[fjutŋi]	[fjuŋti]
[dʒɛts]	[fjus]	[dʒɛtz]	[lʌdz]	[glɛtŋə]	[glɛŋtə]
[fjuts]	[klɛs]	[fjutz]	[swɛlvz]	[kɛtŋi]	[kɛŋti]
[klæts]	[pʌsk]	[klætz]	[θɝz]	[skotŋə]	[skotŋtə]
[pʌts]	[stɪ]	[pʌtz]	[zɛkst]	[slutŋə]	[slutŋtə]
[θʌts]	[svɛt]	[θʌtz]	[zɛlft]	[smɪŋi]	[smɪŋti]
[vɪts]	[swɪ]	[vɪtz]	[zɛmt]	[spitŋə]	[spitŋtə]
[wots]	[tɪmps]	[wotz]	[zʊst]	[jɪŋi]	[jɪŋti]
[yɪts]	[θʌs]	[yɪtz]	[zʊθt]	[zotŋi]	[zotŋti]

The author used a head-mounted microphone in a sound-treated booth to record all the items. The [tz] items (Table 3, column 3) were created by splicing a [z] into the [ts] items. Figure 1 shows the waveforms and spectrograms of a [ts] item [blɛts], and the corresponding spliced [tz] item [blɛtz]. This shows that the only difference between the two sets of items is in the final strident.

Figure 1. Waveforms and Spectrograms of [blɛts] (top) and [blɛtz] (bottom), showing the spliced [z] in [blɛtz].



From figure 1, the reader will notice that there is a difference in length and amplitude in the final stridents. The implications of these differences, and potential problems they create will be discussed below in section 5.

The items in the [ŋ] group differ from the other two groups in that they are all bi-syllabic. The fillers in this group are also bi-syllabic. The [ŋ] always occurs at a syllable

boundary, which places the novel sequence [tŋ] across a syllable boundary (Table 3, column 5). Each item with a [tŋ] sequence is paired with an item that differs only in the positions of [t] and [ŋ], to produce an [ŋt] sequence, also across a syllable boundary (Table 3, column 6). Since these items were produced naturally without any splicing, there are differences in length of the [ŋ] for each item. Besides this possible problem, there is sure to be nasalization on the vowel for the [ŋt] items, which may speed response times.

3.3. Procedure

The experiment was constructed using the E-Prime experimental software. Participants were seated in a sound-treated booth with a computer screen and a button box. They listened to the stimuli played through well-fitting headphones. The experiment consisted of three parts, corresponding to the three stimuli groups. At the beginning of each part of the experiment, participants received written instructions on the computer screen that explained which sound to listen for, and to press the ‘yes’ button as quickly as possible if the item had the target sound, or the ‘no’ button if it did not. Participants were then given five practice items with feedback on whether they responded correctly. After the practice items, participants heard the 50 items from the A, B, and C groups in random order. After these 50 items, participants received the instructions for the next target sound. Half of the participants heard the groups in the order [s], [z], [ŋ], and the other half had the order [s], [ŋ], [z]. The entire procedure lasted about 10 minutes.

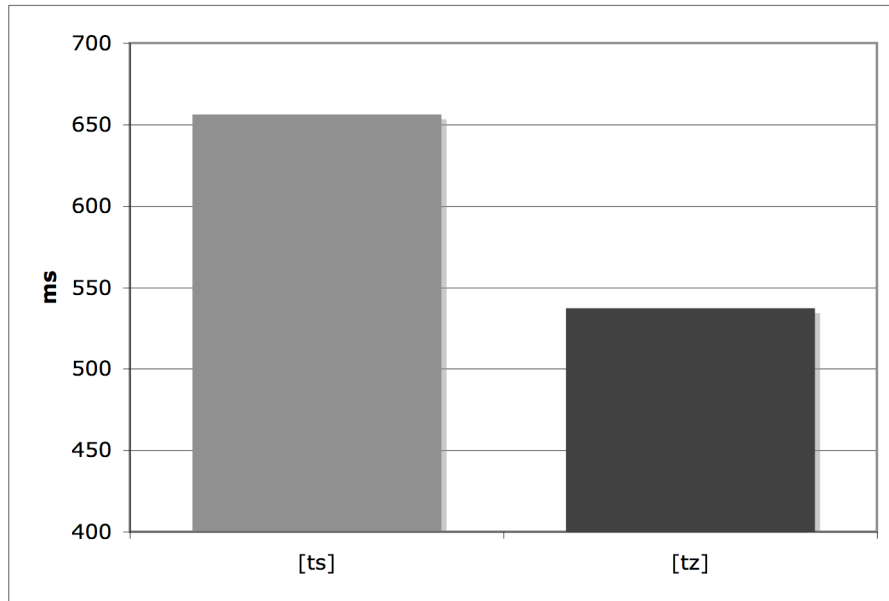
4. Results

All response times were measured from the beginning of the target sound. If the participant answered incorrectly (pressed “no” for an item that contained the target sound), the response was not included in the result. All response times higher than 1500 ms were also excluded from the results. If the participant had an error rate higher than 50% for any part of the experiment, all data was excluded for that part. Similarly, if more than 50% of a participant’s responses were over 1500 ms, that data was also excluded. This led to the exclusion of 2 participants’ data for the [s] portion, and 8 participants’ data for the [ŋ] portion. The high error rate for the [ŋ] portion can be attributed to unclear instructions for that portion for the first 7 participants. These instructions were modified and error rates subsequently fell to tolerable levels. No subjects had high error rates or slow response times for the [z] portion.

4.1. [s] vs. [z]

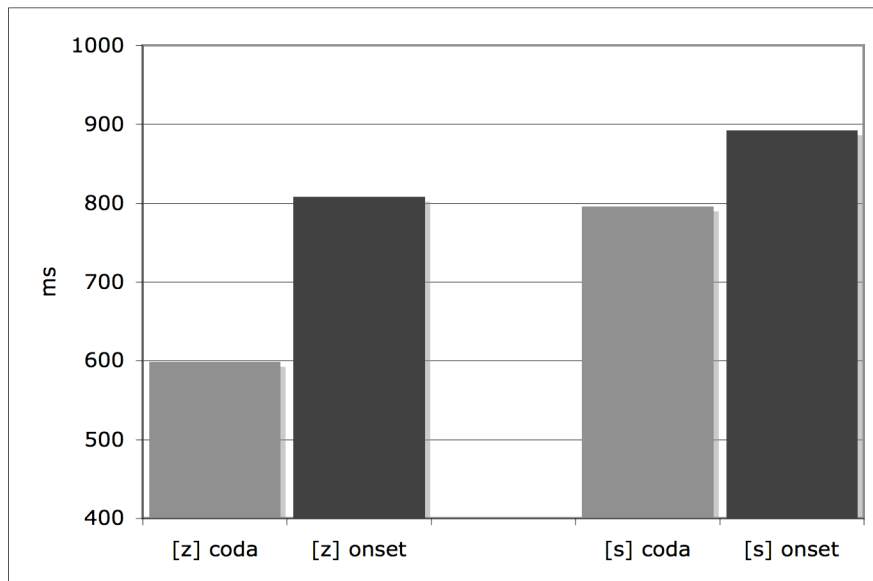
Contrary to predictions, the rule-violating [tz] sequence was responded to 119ms faster than the non-violating [ts] sequence, as Figure 2 shows. Participants were monitoring for [s] in the [ts] items, and for [z] in the [tz] items.

Figure 2. Mean response times for [ts] (656ms) and [tz] (537ms). [tz] is 119ms faster than [ts].



The B items (the items that have [s] and [z] in other environments) also show a difference in response time between [z] and [s]. Figure 3 shows that [z] has faster response times regardless of where it appears in the word, whether it is in the onset or the coda.

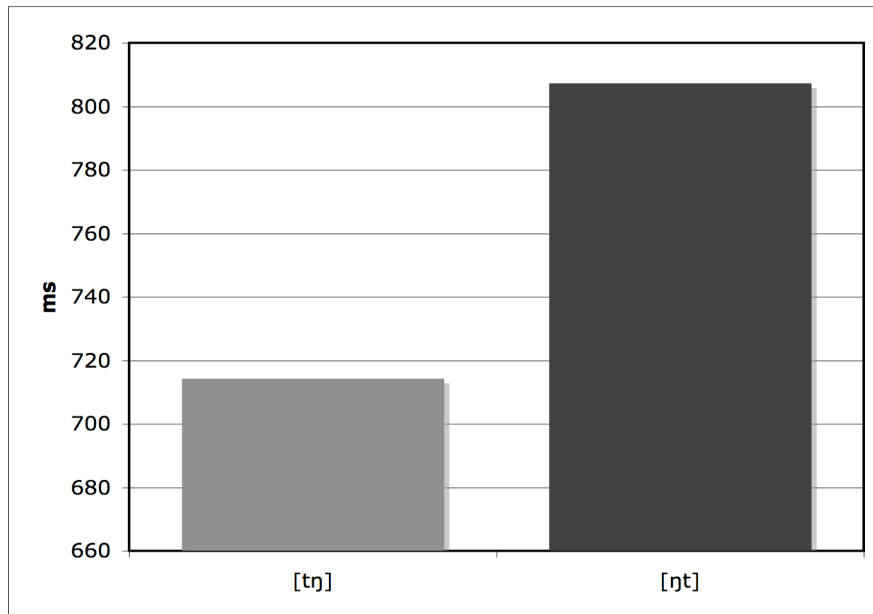
Figure 3. Mean response times of the [s] and [z] items that do not follow [t]. The items are split according to whether they are in the onset or the coda of the word. [z] is faster than [s] in both environments, and target sounds in the coda are faster than in the onset.



4.2. [ŋ]

The novel sequence [tŋ] was responded to faster than the non-novel [ŋt], by an average of 93ms, as shown in Figure 4.

Figure 4. Mean response times for [tŋ] (714ms) and [ŋt] (807ms).



5. Discussion

5.1. [s] vs. [z]

The [tz]~[ts] results in Figure 2 seem to contradict Weber's findings that a violation of a progressive rule must produce a novel sequence in order to speed up response times. There are several possible reasons why the rule-violating [tz] sequence had faster reaction times. First of all, perhaps [tz] really is a novel sequence, even though it occurs orthographically across word boundaries as in the sentence *the cat zipped up the stairs*. Phonetically, the final /t/ in *cat* is almost always realized as an unreleased [t̚] or glottal stop, even in careful speech. Since there is a phonetic distinction between the

sequences [tz] and [t̄z], this may qualify the [tz] sequence as novel, and therefore the results fall in line with Weber's findings.

A second possible explanation for the faster [tz] response times may be found in comparison of the waveforms for [s] and [z] shown in Figure 1, where it can be seen that [z] has a slightly higher amplitude than [s]. This higher amplitude may render the [z] easier to process, and thus explain the faster response times.

It is also possible that [z] is processed faster than [s], regardless of the context, which seems to be supported by the findings presented in Figure 3. Perhaps the phonetic differences and frequency difference between [s] and [z] somehow make [z] naturally easier to process. If this is the case, then the faster response times for the [tz] sequence may be due to these natural differences, rather than to novel popout. A future experiment will explore this possibility in greater depth. Rather than testing [tz] against [ts], this future experiment will test non-rule-violating [dz] against rule-violating [tz]. This experiment will eliminate the effect of phonetic differences between the target phonemes.

5.2. [ŋ]

The [ŋ] results presented in Figure 4 are interesting in that they suggest that a novel sequence is processed faster than a corresponding non-novel sequence, even if there is no assimilation rule involved. Although novel sequences produced by a regressive rule violation are processed slower, as shown in Otake et al. (1996), this inhibition is probably due to the violation of the strong expectations produced by the regressive rule. There is a possibility that the [ŋt] sequence is slower because of a violation of strong expectations, but [ŋt] is not a novel sequence, since it occurs across

word boundaries, as can be seen in the sentence *Sing to me*. Since the [ŋt] sequence in this experiment occurs across a syllable boundary, regressive assimilation violations should not come into play. Needless to say, the findings of this portion of the experiment deserve further exploration in future studies to see if other novel sequences facilitate processing independently of other factors, as seems to be the case in the present study.

5.3. Future experiments

Besides the future experiment discussed above in section 5.1, I believe it would be worthwhile to investigate the phenomenon of novel popout in other phonological dimensions, for example in tone. Mandarin Chinese offers an interesting possibility. Mandarin is traditionally described as have four contrastive tones, along with a fifth neutral tone. This neutral tone is realized at different pitches, depending on the tone of the preceding word. It would be interesting to violate this progressive rule to produce novel tone-pitch sequences and see whether this resulted in a novel popout effect. It would also be interesting to look for progressive assimilation rules in other languages, which could be manipulated in a similar manner to test for novel popout.

References

- Gaskell, M., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 144-158.
- Gaskell, M., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 380-396.
- Johnston, W., & Schwarting, I. (1996). Reassessing the evidence for novel popout. *Journal of Experimental Psychology: General*, **125**, 208-212.
- Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. H. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustic Society of America*, **100**, 3831-3842.
- Weber, A., (2001). *Language-specific listening: the case of phonetic sequences*. MPI Series in Psycholinguistics **16**.