SPOKEN WORD RECOGNITION AND ACOUSTIC CUES OF FINAL VOICELESS STOPS*

Mercedeh Mohaghegh
University of Toronto

Participants followed spoken instructions about objects whose name ended in either a released or unreleased voiceless stop. The status of background noise was also varied from absent to present across trials. The results indicated less success in identification of the intended word when the final sound of a target word (e.g. cat) was unreleased, as reflected by increased confusion with a competitor word (e.g. cap). No effect of background noise was found. Examination of on-line eye movement data, on the other hand, revealed a rapid influence of the acoustic cues provided by vowel transitions on the activation of the target noun. However, these measures revealed no such sensitivities in noise, suggesting listeners' judgment scores were more likely to be based on mere chance or a more general "off-line" impression of the sound pattern of a word.

1. Introduction

As listeners hear words unfold in time, information carried by individual speech sounds is continuously mapped to stored lexical candidates until only a single candidate remains. However, the identification of individual speech sounds, as required by the word recognition process, is itself of a complicated nature. Variation in the form of speech signals as they are produced and additional background noise are among the factors that contribute to these complications. Nevertheless, spoken word recognition is considered to be very robust. How is the robustness of lexical processing achieved and affected by factors such as noise and acoustic variation under different conditions? As a means to investigate this question further, the current research examines the case of English released and unreleased voiceless stops in word/utterance-final position and the effect of noise on their perception.

Acoustic-phonetic investigations have revealed that the release portion of stops carries important acoustic cues as to their place of articulation (PoA). However, the production of the release, regardless of whether the stop occurs in prepausal position or word internally in the middle or at the end of an utterance, is found to be optional at times (Crystal & House, 1988; Davidson, 2011; Lisker, 1999; LoCasto & Connine, 2011). For example, in an analysis of 1,130 word final stops by Byrd (1993), 40% of them were found to be produced with no final release. According to Crystal and House (1988), voiced stops are more prone to be unreleased than voiceless stops (82% vs. 58%). In addition, velar stops are less often produced with inaudible release (17%) compared to coronals (43%) and labials (50%; Byrd, 1993). Other studies have also shown that stops

* I am most grateful to Dr. Craig Chambers, Dr. Yoonjung Kang and Dr. Elizabeth Johnson for providing much help and support with their useful insights on this topic.

Actes du congrès annuel de l'Association canadienne de linguistique 2012.
Proceedings of the 2012 annual conference of the Canadian Linguistic Association.

tend to be released more overall when the preceding vowel is tense than when it is lax (Kang, 2004; Parker & Walsh, 1981). Other factors, such as sociolinguistic factors relating to speech style, the gender of the speaker, and social status, are also noted to have an effect on choosing between the two variants (Davidson, 2011; Deelman & Connine, 2001; Lisker, 1999). Lisker (1999) conducted an experiment on the perception of final [p˺,t˺,k˺] preceded by 14 different vowels (7 monophthongs and 7 non-monophthongs/diphthongs) in 84 nonsense syllables. He used a forced choice judgement task that required participants to choose among /p/, /t/ and /k/ when pairs of stimuli were presented through headphones. The results indicated a surprisingly high rate of correct responses in the unreleased condition (almost 90%). Lisker also found an effect related to the nature of the preceding vowel, with non-monophthongs resulting in overall lower intelligibility of the final stops compared to monophthongs, particularly in case of /k/. He also examined the effect of naturally produced unreleased stops on perception as opposed to dereleased ones, which were engineered by removing the release (and aspiration) portion of the sound through editing. He found no difference between these two conditions. A comparison of identification scores based on different places of articulation in Lisker's study revealed an overall higher rate of correct responses for labial stops, followed by coronals and finally velars.

When information carried by the stop release burst is eliminated, the vowel transitions (the transition of F2 from the preceding vowel and in some cases F3) remain as the available cues to PoA (Delattre, 1958; Halle et al., 1957; Hume et al., 1999). This brings us to the question of how processes above the level of segment perception are affected by subphonemic details, such as vowel transition and stop release information. Previous studies on spoken word recognition provide evidence highlighting listeners' sensitivity to subphonemic coarticulatory effects as speech unfolds in time (Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Marslen-Wilson & Warren, 1987; McQueen, Norris, & Cutler, 1999). Dahan et al. (2001) created cross-spliced stimuli by excising the final stop from two CVC real words, the target (e.g., *net*) and a "competitor" (e.g., *neck*), as well as a corresponding nonword (e.g., *nep*). They then added the consonant from the final position of the target word (e.g., -/t/) to the three CV-stems, creating three versions of the target word stimulus. As participants heard words presented to them in each of the three splicing conditions within an instruction (e.g., "now the net"), their eye fixations to pictures associated with the target words were monitored. The remaining three pictures in the display were either all phonologically-unrelated distractors (Exp. 1), or were the real-word competitor (*neck*) and two distractors (Exp. 2). The results of the first experiment showed slower identification of target pictures when the coarticulatory cues in the target word were drawn from a different (real) word than the naturally produced target word form (e.g., when the vowel came from *neck*), compared to when they corresponded to a nonword (e.g., when the vowel came from *nep*). This indicated lexical competition under the former condition but not the latter one. In the second experiment, when the mismatching acoustic cue was drawn from the real word competitor, comparatively more temporary fixations to the competitor picture were observed, further confirming sensitivity to the subphonemic acoustic details (i.e., mismatching vowel transitions) over time. These results support the view that listeners take advantage of subphonemic details when they become available to them and that the spoken word recognition system updates itself in time as it is continuously provided with successive acoustic details.

In the current study, the influence of subphonemic cues on the activation of lexical items was investigated using released and unreleased final voiceless stops p, t, k when background noise was present or absent. Word identification in noise can provide information about degree of robustness of the word recognition system when the relevant acoustic information is degraded. In this study, broadband speech-spectrum noise with a signal-to-noise ratio of +4 dB SPL was added. This type of background noise makes listening more challenging by interfering with speech but does not selectively mask information carried by the sounds. The process of spoken word recognition was investigated using an eye movement paradigm. This method was chosen to examine dynamic aspects of lexical activation because eye fixations to relevant pictures have been shown to be more sensitive to fine-grained acoustic cues as speech unfolds over time compared to offline techniques such as a phoneme monitoring task (Dahan & Gaskell, 2007; McMurray et al., 2008; Tanenhaus, Magnuson, Dahan, & Chambers, 2000).

2. Method

2.1 Participants

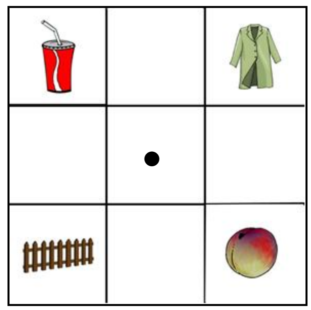
Twenty four adult English native speakers within the age range of 17-27 years (M = 18.7, SD = 2.29, 8 males and 16 females) participated in this study. All participants were native speakers of English. None of the participants reported any visual or auditory problems.

2.2 Materials

For a given trial, the visual stimuli consisted of four clip-art pictures presented in the corners of a 3 x 3 grid on a computer screen (see example in Figure 1). On critical trials, each display included a target image (e.g., [kowk] ‘Coke’) and a "competitor", whose name was a minimal pair for the target item, phonetically similar in all sounds except for the last consonant (e.g., [kowt] ‘coat’). The last consonants in both the target and competitor names were always voiceless stops /p/, /t/ or /k/ (e.g., [ɹæk] ‘rack’ and [ɹæt] ‘rat’), and were preceded by a vowel. The two remaining "distractor" pictures were neither phonologically nor semantically related to the target and competitor names (e.g., [fɛns] ‘fence’ and [pitʃ] ‘peach’ being distractors in a set with [wɪp] ‘whip’ and [wɪk] ‘wick’). The positioning of the four images in each display was randomized across trials.

The auditory stimuli accompanying the displays were recorded from a female Canadian English native speaker (Toronto accent). On critical trials, the recorded materials consisted of an initial instruction ‘Where is the X?’, where X was replaced by the target word, and a second instruction ‘Now, where is the Y?’, where Y was replaced by the name of one of the distractor pictures. The recorded target-competitor stimuli consisted of sixteen pairs of nouns. In addition, 16 words referring to the picture distractors were recorded, all of which were monosyllabic nouns to be presented in the critical displays. Following Lisker (1999), the vowels preceding each of the final voiceless stops were categorized either as a monophthong /ɪ æ ɛ ʌ ɑ/ or a diphthong (non-monophthong) /ij uw ej ow/. As a result, the syllabic nucleus in 11 out of 16

Figure 1. Example 3 x 3 grid used on a critical trial. In this example, two of the images correspond to target-competitor nouns (Coke, coat) and the other two images are distractors (fence, peach). The small black circle in the middle of the grid disappeared before the auditory and visual stimuli were presented.



pairs was a monophthong and in the remaining five pairs was a diphthong. Each target word on critical trials occurred in one of the two versions corresponding to released and unreleased experimental conditions. In order to ensure the acoustic comparability of the items across these conditions, the version in which the final consonant in the target name was unreleased was created by manually cutting the release (and aspiration) portion of the stop in the original recording (produced with the release) using Praat (Boersma & Weenink, 2010). A period of silence of the same duration as the removed portion of the original final consonant was then added. In order to create stimuli for the noise condition, the stimuli were then mixed with speech spectrum noise with a signal-to-noise ratio of +4 dB SPL.

The released/unreleased final consonant manipulation was crossed with the noisy/quiet background condition, creating four main experimental conditions (released-quiet, released-noisy, unreleased-quiet, unreleased-noisy). For the analysis of eye-movement data, only the unreleased-quiet and unreleased-noisy conditions were of interest. These conditions can reveal the role of the vowel cues on real-time word recognition processes when the release cues to PoA are absent.

In addition to the critical displays, 16 filler displays were included in order to divert the participants' attention from the main goal of the experiment. The four clip art pictures used in filler displays were also associated with monosyllabic nouns. Fillers and critical displays were randomly intermixed. However, the trials were presented sequentially in two blocks so that the first block (8 critical and 8 filler trials) was presented without background noise whereas the second set (8 critical and 8 filler trials) was presented with background noise.

2.3 Procedure

Participants were tested individually. The visual stimuli were presented via a computer screen, placed on a table. Participants were seated at a comfortable distance from the computer screen, and auditory stimuli were presented through stereo speakers located 0.5 meters from participants' seat. A head-mounted eyetracker, which sampled eye positions every 2 ms, was used to record the eye movements and gaze fixation points throughout the experiment.

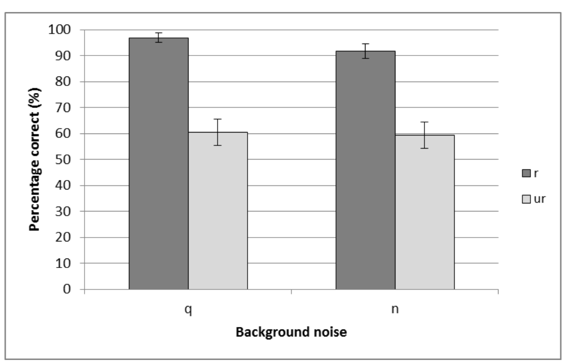
Each experiment began with some general instructions provided by the experimenter and a calibration procedure. After the calibration process, the experimental trials were initiated by the experimenter by pressing a keyboard button. The first two trials were for practice and closely resembled the experimental and filler trials in terms of their structure. On each trial, a set of four clip art images appeared on the screen on a 3 x 3 grid and the first auditory instruction began approximately 500 ms later. Participants were instructed to freely look at the images on the screen and to respond to the questions by saying out loud where the target image was located on the grid (e.g., top-left, bottom-right, etc.).

3. Results and discussion

3.1 Identification scores

The data were analyzed for the accuracy of participants' final responses in selecting the intended picture after hearing the first instruction, 'Where is the X?'. The percentage of correct responses based on the released/unreleased and quiet/noise conditions is depicted in Figure 2. In this figure, the percentage of correct identifications in the released condition is clearly higher than in the unreleased condition in both quiet and noise. However, background noise did not lower the accuracy responses to a great extent (maximum 5.2% lower in released and 1.1% in unreleased). A repeated measures analysis of variance (ANOVA) was conducted on the average correct responses, with release (released, unreleased) and noise (quiet, noise) as within subject factors. The results showed a highly significant main effect of release, $F(1,23) = 86.07$, $p<.001$, $\eta^2=.79$, which reflects higher accuracy rates in released condition, as would be expected. The lack of release results in a decrease of around 30-40% in accuracy rates (Figure 2). No reliable effect of the noise manipulation was detected, nor was the interaction of noise and release significant.

Figure 2. Percentage of correct word identifications across release (r), unreleased (ur), quiet (q) and noise (n) conditions. The error bars represent the amount of standard error.

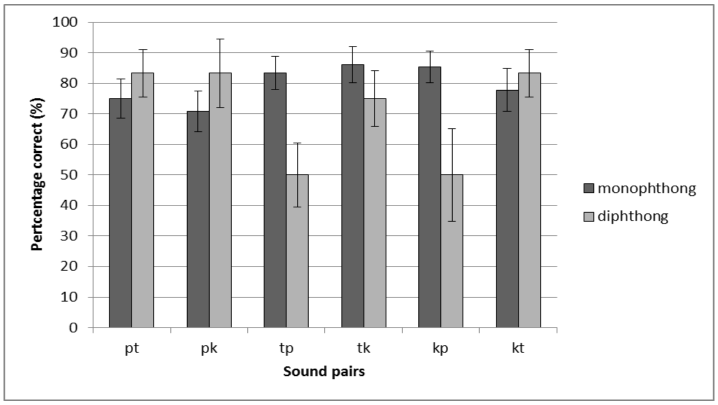


	Target C	Released		Unreleased	
		Mono	**Diph**	**Mono**	**Diph**
Quiet	**p**	91.7	100	50	66.7
	t	100	91.7	71.4	41.7
	k	100	100	81	33.3
Total		97	96.7	66.7	46.7
Noise	**p**	79.2	88.9	70.8	77.8
	t	100	83.3	66.7	33.3
	k	100	100	47.6	55.6
Total		92.4	90	62.1	53.3

Table 1. Mean percentage of correct responses in released and unreleased conditions with no background noise. The data are grouped according to the two vowel types (monophthong and diphthong) and the three final consonants in the target word.

In view of Lisker's (1999) results, the identification scores were also analyzed based on the type of vowel preceding the final stops. Table 1 shows that, in the released condition, words tended to be recognized accurately regardless of the type of vowel. However, in the unreleased condition, final stops that followed a monophthong (e.g., /t/ pronounced after /æ/ in 'cat' [kæt]) tend to show an overall higher accuracy score compared to those following a diphthong (e.g., /t/ pronounced after /ow/ in 'coat' [kowt]). In order to test the significance of the effects of vowel type and final consonant, a separate mixed-model ANOVA was conducted on average correct responses across items with final consonant release (released, unreleased) and acoustic background (quiet, noise) as within-item factors and vowel type (monophthong, diphthong) and the final consonants in the target-competitor pairs (pt, pk, tp, tk, kp, kt) as between-item factors. The results of ANOVA revealed no main effect of any single factor. However, the interaction between vowel type and consonant pair turned out to be significant, $F(5, 20) = 2.7$, $p = .05$, $\eta^2 = .32$. Paired contrasts revealed a significant difference between average correct responses in context of monophthongs ($M = .84$, $SD = .1$) versus diphthongs ($M = .5$, $SD = .1$) when the final consonant in the competitor was /p/, t (9), $p = .001$. These results indicate that the final stops /t/ and /k/ are overall more prone to be confused with /p/ in the context of a diphthong vowel (this interaction can be seen more clearly in Figure 3, which collapses across the two acoustic background and release conditions).

Figure 3. Percentage of correct responses across vowel type (monophthong vs. diphthong) according to the final consonants in the target name (first consonant labeled in each category) vs. the competitor name (the second consonant labeled). The error bars represent the standard error.



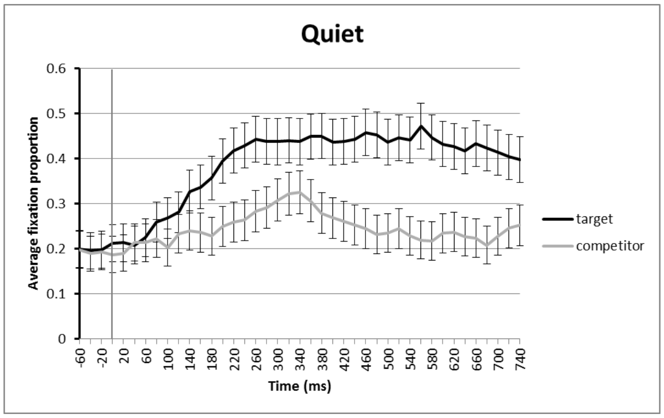
3.2 Eye movement data

In order to investigate how vowel transition cues influence the perception of the PoA of final stops when the final stop is not fully articulated, the eye movement data were analyzed in the unreleased consonant conditions. These data can reveal the time course of listeners' sensitivity to vowel cues as the vowel portion of the target word becomes available. The eye fixations were recorded by the tracking system beginning at the onset of the target noun. The proportion of fixations over time to the target, competitor and the two unrelated pictures was calculated. The fixations were aligned at onset of stop closure in the target word to facilitate comparisons across items and conditions for the point at which information from the preceding vowel might drive eye movements to the target or competitor objects. The transition of the vowel to final consonants was assumed to occur shortly (approximately 60 ms) prior to the onset of stop closure.

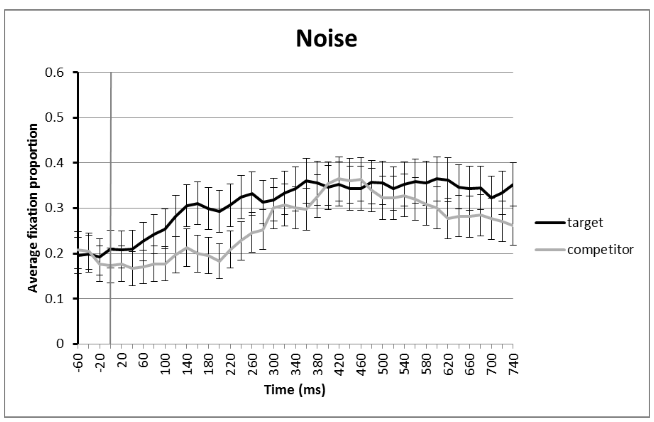
The data were analyzed for effects of vowel transition and noise on eye fixations. Figures 4a and 4b depict the patterns of fixations on target and competitor pictures in the quiet and noise conditions respectively. Inspection of Figure 4a shows that fixations on pictures corresponding to target nouns started to diverge from fixations on the pictures corresponding to competitor nouns on average 60 ms after the onset of stop closure. In studies of this type using eye movement measures, a delay of 100-200 ms is considered to reflect the latency in initiating and launching an eye fixation in response to speech input (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Magnuson et al., 2001). Because vowel transitions were expected to occur within 60 ms prior to onset of stop closure the additional launch latency would suggest that increasing fixations on the target at around 60 ms are in fact based on the point in the speech stream when vowel transition cue became available. However, in the presence of background noise, the point of divergence is not clear (Figure 4b). A short-lived

Figure 4. Mean fixation proportion on images associated with target (black line) and competitor (grey line) nouns over time.The zero mark on the horizontal axis (grey vertical line) corresponds to the onset of stop closure. The vowel transition region is within the 60 ms prior to the zero point on the time axis. Panel 4a depicts fixations in the unreleased-quiet condition and fixations in unreleased-noise condition are depicted in Panel 4b.

4a.



4b.



advantage in favor of target pictures can be observed around 40 ms, which is again close to the same time window where the vowel transition cue became available; however, the overall pattern of eye movements does not indicate a sustained advantage of target nouns over competitor nouns when there is noise in the background environment. This pattern might suggest that the noise did not fully mask the vowel cues, but has nonetheless introduced uncertainty into the recognition system in other ways. This uncertainty might reduce the importance of secondary cues even if they are perceived at some level.

In order to test the statistical significance of the observed patterns, a within-subjects repeated-measures ANOVA was conducted. To ensure that the measures are sufficiently sensitive, a frequency measure was included in the analysis. Previous studies have indicated highly consistent effects of words frequency on lexical processing, with higher-frequency words recognized more quickly than lower-frequency ones. Thus, a failure to find any effects, including frequency effects, would indicate a lack of measurement sensitivity. The factors included in the ANOVA were frequency (high, low) as well as noise (quiet, noise). The test was conducted using a single "target advantage" measure that captures listeners' unconscious ability to discriminate the target word from the name of the competitor object. This was achieved by calculating the difference in fixation proportions to the target minus the competitor averaged over time (-60 to 740 ms relative to onset of stop closure). The results indicated that in fact there is a significant effect of word frequency on the patterns of eye fixation, F (1, 23) = 10, p = .004, η^2 = .30. Somewhat surprisingly, neither an effect of noise nor an interaction of noise and frequency was found, despite the pattern observed in Figure 4. However, conducting the same test of ANOVA for a more limited time window (340-740 ms relative to onset of stop closure), reflecting the latter phases of lexical processing, revealed a marginally significant effect for the noise factor as well, F (1, 23) = 3.1, p = .09, η^2 = .03.

4. Conclusion

In the current experiment word recognition scores and patterns of eye-movement were analyzed. The results indicated that the release of the final consonant plays an important role in correct identification of lexical items. However, the presence or absence of background noise with an SNR of +4 dB SPL was not found to highly influence recognition scores. The analysis of eye-movement data on the other hand indicated a rapid influence of the acoustic cues provided by vowel transitions on the activation of the target noun right at the earliest point this information becomes available to the listeners. This is consistent with previous eye-tracking studies on time course of activation of lexical items (e.g., McMurray et al., 2008), which support the idea that word recognition involves integration of cues as soon as they become available. However, unlike final accuracy scores which were comparable between the two noise conditions, in the unreleased-noise condition listeners did not benefit from vowel transition cues as much as they did in the unreleased-quiet condition. This might suggest that, in the noise condition, participants' judgements (on average 59% correct) were more likely to be based on mere chance. Another possibility is that, when listening in noise, perceivers may base their decisions on a more general "off-line" impression of the sound pattern of a word, rather than using the information in a more linear way as the sound pattern unfolds in time.

The current data do not reflect the specific effects of vowel type that were noted previously (e.g., those reported by Lisker, 1999). Although diphthongs were found to increase the likelihood of misidentifying the target word as having a final stop with labial place of articulation, there was no overall benefit for perception of place of articulation of stops when the preceding vowel was a monophthong as opposed to a diphthong. However, the differences between the designs of the current and previous experiments might explain the differences in the final results.

In future investigations, it is possible to present the stimuli using their orthographic representation rather than clip art images, making it possible to include words that are difficult to depict as clip art objects. It will also be of interest to further explore the role of subphonemic cues on listeners' performance using other types of stimuli such as words containing voiced final stops, as well as controlling for factors such as vowel context, listeners' native language and their knowledge of other languages besides their native tongue.

References

Abramson, A. & Tingsabadh, K. 2000. Thai final stops: cross-language perception, *Phonetica*, *56*, 111-122.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Ben-David, B., Chambers, C., Daneman, M., Pichora-Fuller, M., Reingold, E., & Schneider, B. 2011. Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research, 54,* 243-262.

Byrd, D. 1993. 54,000 American stops. *UCLA Working papers Phonet., 83*, 1-19.

Crystal, T. & House, A. 1988. The duration of American-English stop consonants: An overview. *Journal of Phonetics, 16*, 285-294.

Dahan, D., Magnuson, J., Tanenhaus, M., & Hogan, E. 2001. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16,* 507–534.

Dahan, D., & Gaskell, M. 2007. The temporal dynamics of ambiguity resolution: Evidence from spoken-word recognition. *Journal of Memory and Language, 57,* 483–501.

Davidson, L. 2011. Characteristics of stop releases in American English spontaneous speech. *Speech Communication, 53*, 1042-1058.

Deelman, T. & Connine, C. 2001. Missing information in spoken word recognition: nonreleased stop consonants. *J. Exp. Psycholing.: Human Percept. Perform. 27*, 656–663.

Delattre, P. 1958. Unreleased velar plosives after back-rounded vowels. *Journal of the Acoustical Society of America, 30,* 581-582.

Dinnsen, D., & Charles-Luce, J. 1984. Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics 12*, 49–60.

Garellek, M. 2011. The benefits of vowel laryngealization on the perception of coda stops in English. UCLA Working Papers in Phonetics, *109*, 31–39.

Halle, M., Hughes, W. & Radley, J. 1957. Acoustic properties of stop consonants. *JASA, 29*, 107–116.

Householder, F. 1956. Unreleased PTK in American English. In Halle, M., Lunt, H., McLean, H. & van Schooneveld, C. (Eds.) *For Roman Jakobson: Essays on the occasion of his sixtieth birthday*. The Hague: Mouton, 235–244.

Hume, E., Johnson, K., Seo, M., & Tserdanelis, G. 1999. A cross-linguistic study of stop place perception. *Proceedings of the 14th International Congress of Phonetic Sciences.* 2069-2072.

Jun, J. 1995. Place assimilation as the result of conflicting perceptual and articulatory constraints. *West Coast Conference on Formal Linguistics, 14*, 221-237.

Kang, Y. 2004. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology, 20*, 173–218.

Kewley-Port, D., Pisoni, D., & Studdert-Kennedy, M. 1983. Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America*, *73*, 1779-1793.

LoCasto, P. & Connine, C. 2011. Processing of no-release variants in connected speech. *Language and Speech, 54*, 181-197.

Lisker, L. 1999. Perceiving final voiceless stops without release: effects of preceding monophthongs versus nonmonophthongs. *Phonetica, 56*, 44-55.

Malecot, A. 1958. The role of release in the identification of released final stops: a series of tape-cutting experiments. *Language, 34*, 370-380.

McMurray, B., Clayards, M., Tanenhaus, M., & Aslin, R. 2008(b). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin and Review, 15,* 1064-1071.

McMurray, B., Tanenhaus, M., & Aslin, R. 2002. Gradient effects of within-category phonetic variation on lexical access, *Cognition, 86,* B33-B42.

McQueen, J., Norris, D., & Cutler, A. 1999. Lexical influence in phonetic decision making: evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1363–1389.

Ohde, R. N., & Sharf, D. 1981. Stop identification from vocalic transition plus vowel segments of CV and VC syllables: A follow-up study. *Journal of the Acoustical Society of America, 69*, 297-300.

Parker, Frank & Thomas Walsh. 1981. Voicing cues as a function of the tense/lax distinction in vowels. *Journal of Phonetics*, *9*, 353–358.

Port, R. & O'Dell, M. 1985. Neutralization of Syllable-Final Voicing in German. *Journal of Phonetics 13*, 455-472.

Rossion, B., & Pourtois, G. 2004. Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic level object recognition. *Perception*, *33*, 217–236.

Stevens, K., & Blumstein, S. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, *64*, 1358-1368.

Tanenhaus, M., Magnuson, J., Dahan, D. & Chambers, C. 2000. Eye movements and lexical access in spoken-language comprehension: evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*, 557–580.

Warren, P., & Marslen-Wilson, W. 1987. Continuous uptake of acoustic cues in spoken word-recognition. *Perception & Psychophysics, 41,* 262-275.