

Orthographic effects on the attention to phonetic cues¹

Grant McGuire

Department of Linguistics

University of California at Santa Cruz

Stevenson Faculty Services

Santa Cruz, CA, USA 95064

Email: gmcguir1@ucsc.edu

Telephone: 831-459-4988

Abstract: This paper presents experimental data that attempts to tie together two different phenomena. First, there is strong evidence that phonemic awareness is strongly influenced by one's writing system such that learning an alphabetic writing system focuses attention on the segment as a unit separate from other larger units, such as syllables. This phenomena may be related to the Developmental Weight Shift hypothesis (Nittrouer, 1996) where English learning children switch from relying on dynamic formant transitions to static frication noise in sibilant identification. This could be accounted for as a shift to phonemic processing brought on by learning to read and write English. This paper presents data from experiments with adult native-speakers of English demonstrating that attention to sibilant cues can be directed orthographically to attend to different cues, depending on the symbol set used. The ramifications of sub-segmental orthographic effects on theories of speech perception and subject population is discussed.

¹ Thanks go to Meghan Sumner for initially suggesting I follow up on the orthographic possibilities initially found in my dissertation. Thanks also to the Phonology Laboratory at UC Berkeley for access to their facilities for part of this research. Thanks also go to Angela Aiello, Devin Tankersley, and Lauren Winens for assistance with subject running. A portion of this research was funded by Keith Johnson's NIDCD Grant #R01 DC004421.

1. Introduction

There is considerable evidence that phonemic awareness (PA), the ability to decompose the speech signal into segment sized units, is strongly influenced by one's writing system, such that learning an alphabetic writing system focuses attention on the segment as a unit separate from other larger units such as syllables (e.g. Cheung et al. 2001). However, the articulations that give rise to the acoustic phonetic cues to such contrasts are gradual and do not neatly fit into isolated segment gestures. The consequence of these coarticulatory effects is that the perceptual cues for identifying a given contrast may be located not only in the segment itself, but also the adjacent segments. For some contrasts, such as plosives, these “adjacent” cues are in fact the most salient (Rafael 2005). This paper demonstrates that orthography can impact such sub-segmental cues and provides evidence that the sibilant cue weighting shift seen in English acquiring children (Nittrouer and Studdert-Kennedy 1987) may be due to learning to read and write.

The primary evidence for orthographic influences on speech perception comes from three sources: the correlation between PA and reading ability in children, comparisons of literate and non-literate adults, and comparisons of speakers having different orthographic traditions. As to the first, there is a long history of evidence that children with better reading abilities also show greater PA (see Bentin 1992 for an overview). While pre-literate children are aware of onset vs. rime characteristics (Kirtley et al. 1989), they are generally unable to fully decompose a word into phones (Bruce 1964). However, children with more schooling show improved abilities at PA, independent of age (Bentin et al. 1991), and children who are better readers overall compared to their peers also perform better at PA tasks (Lyon 1998).

As might be expected given the acquisition data, literacy and PA effects are also found in adults. In an important series of studies, Morais and colleagues found that Portuguese illiterates performed much more poorly than literates on a series of PA tests (Morais et al. 1979, 1996). These studies demonstrated that illiterate adults had extreme difficulty in identifying onsets, unlike the literate subjects. While the illiterates were capable of identifying syllable and rime components, the literate subjects were considerably better at that task, too. Ex-illiterates showed intermediate abilities.

All of the research above deals with subjects who speak (or are acquiring) a language that uses an alphabetic orthography; other evidence shows that non-alphabetic systems do not help with PA. For example, Read et al. (1986) found that Chinese speaking adults literate only in Chinese characters (logographic) performed similarly to the illiterates in Morais et al. (1979), while those with training in Pinyin (alphabetic) performed similarly to the literates. Such effects are robust and far-reaching, for instance Cheung and Chen (2004) found that these orthographic differences not only affect metalinguistic analysis, but also online speech processing using a primed shadowing task.

While the above literature strongly supports the view that many aspects of PA derive from learning an

alphabetic orthography the possibility that sub-segmental cues can be manipulated has been largely left unexplored. Nittrouer (1996) demonstrated that children with higher PA also show more adult-like cue weighting for sibilant fricatives. Sibilants have cues to their place of articulation both in their frication noise, arising from the constriction made by the tongue in the vocal tract as well as cues in the vocalic transitions into and out of the fricative (Johnson, 2003). As Nittrouer and her colleagues demonstrated in earlier work (see e.g. Nittrouer and Studdert-Kennedy 1987, Nittrouer 1992, Nittrouer and Miller 1996), young children tend to weight the formant transitions more than the frication noise, yet, later in childhood, they begin to place more weight on frication noise cues, in much the same way that adults do. While Nittrouer (1996) makes the connection with PA, she does not seem to view orthography as causal, rather the development of PA allows the listener/reader to determine the proper structure.

Given the evidence presented the development of PA, I believe this is not the case. While children seem to be aware of syllable level information from a very early age, the more detailed knowledge required for segmentation only comes after learning to read using an alphabetic system. Therefore it seems quite possible that learning to read and learning that the frication noise corresponds to the <s> and <sh> orthographic symbols is what drives the shift from formant transitions to frication noise. The following experiment is designed to test this hypothesis by using orthography to change the sibilant cue weighting strategies of listeners.

2. Experiments

The following experiments have the same basic design. Listeners first participate in a *same-different* discrimination task to assess how well they can identify the cues involved in the Polish post alveolar sibilant contrast, retroflex /ʂ/ and alveopalatal /ʧ/.² Subjects then are asked to classify the sounds using two labels. There are three groups, each were given a different label set: vocalic cues <sha> / <shya>, frication cues <śa> / <ça>, and both <sha>/<xia>. The first set of orthographic symbols is derived from and experiment where listeners were example sounds from the contrast and asked to type in on a computer what they heard (McGuire 2007). The second set was designed to be plausibly foreign symbols for a sibilant contrast the relied solely on a single symbol changing (i.e. no digraphs). The final set is the Pinyin symbol set for a very similar contrast found in Mandarin Chinese. It was chosen as it shows both a consonantal and vocalic change.

If the orthographic representations do in fact have an effect on cue weighting, then each group should show a different weighting scheme. The vocalic orthography group should preferentially weight the vocalic cues while the frication group should show the opposite effect. The pinyin group, where both symbols

² This task was included for a different purpose (see McGuire 2007) with the first group run. For consistency and the fact that it may have an effect on the classification portion, it was included for all groups.

change, should show an effect on both cues.

2.1. Participants

Three groups of native listeners of English participated. The first group (n = 21) is the *vocalic* group and were recruited from the UC Berkeley campus and paid \$10 to participate³. The second group (n = 18) is the *frication* group and was recruited at the UC Santa Cruz campus and received course credit. The final group (n = 24) is the *multidimensional* group and was also recruited at UC Santa Cruz, receiving course credit. None had any experience with Mandarin Chinese or Polish, though some did have second language experience, primarily Spanish.

2.2. Stimuli

The stimuli used were pulled from McGuire 2007. They consisted of a two-dimensional, 10-step continuum from [ʂa] to [ʐa] varying in frication noise and formant transitions. This continuum was made by taking a representative sample of each sound (produced by native speaker of Polish) in the context of /a/ and separating them at the CV boundary. The frication portions and vocalic portions were separately interpolated to produce two ten-step continua from [ʐ] to [ʂ] (labeled in figures as f0-f9) and [a] excised from [ʐ] to [a] excised from [ʂ] (labeled in figures as v0-v9). These were then recombined into every possible CV combination resulting in 100 stimuli. Perceptual tests with Polish and Mandarin listeners found them to be sufficiently natural as to be representatives of native categories (McGuire 2007).

2.3. Procedure

In the initial *same-different* (AX) discrimination task listeners only heard the four stimuli that represented the endpoints of the continua, i.e. alveolopalatal fricative + alveolopalatal vowel, retroflex fricative + retroflex vowel, alveolopalatal fricative + retroflex vowel, and retroflex fricative + alveolopalatal vowel. These were presented in pairs with a 100ms inter-stimulus interval. There were six possible *different* pairs (four differing in one dimension and two in both dimensions) and four *same* pairs, or repetitions of the four stimuli. There were 96 *different* trials (12 pairs x 2 orders x 8 repetitions) balanced by 96 *same* trials (24 repetitions of each stimulus). The 192 trials were divided into four blocks of 48 trials each with a self-monitored break between them. Subjects had 1500ms to respond and received accuracy and reaction time feedback after each trial.

In the classification portion of the experiment each of the 100 stimuli were presented to the listeners in a randomized order, one per trial. Subjects were given 1500ms to classify each stimulus by pressing the proper button on a labeled box. Subjects were given feedback as to their reaction time. There were 5 blocks

³ Results from these subjects were part of McGuire 2007.

of the stimulus set for a total of 500 trials. The procedure only varied in the labeling of the buttons; the vocalic group was given the labels <sha> / <shya>, the frication group, <śa> / <ça>, and the multidimensional group <sha>/<xia>. All listeners were told that they were hearing sounds from a different language and that they should classify them as best they can. The experiment took approximately 45 minutes to complete.

2.4. Results

In the interest of space, and as it was not the focus of this current experiment, the discrimination portion of the experiment will not be analyzed in detail. For these results the proportion correct for each subject was converted into d' , a measure of sensitivity (MacMillan and Creelman, 2003). The d' values for each orthography group were compared using a single level ANOVA having the factor Group. This factor was not significant ($F[2,195] = 0.60, p = 0.55$), indicating that the groups did not have different sensitivities to the stimuli.

Before the analysis of the classification results, the data from that part were adjusted to account for variation in label assignment. Because listeners could plausibly assign the <śa> label, for example, to either the /ś/ or /ç/ stimuli, the following adjustment procedure was devised. First, each subjects mean labeling response to the [śa] endpoint stimulus was calculated. If more than half were labeled as <sha> for the vocalic and pinyin groups, or as <śa> for the frication group, then no changes were made. If fewer than half were labeled as such, then all the responses were flipped for that subject. This was necessary for 6/18 of the frication listeners and 7/24 of the pinyin listeners. None of the vocalic group participants needed this adjustment.

The adjusted responses were then put in to logistic regression model with the factors fStep, vStep and Group, where fStep refers to fricative step, vStep refers to the vowel step, and the Group factor refers to the orthographic label assignment. The model resulted in significant main effects for all factors and significant interactions between fStep and Group and vStep and Group (see Table 1 for details).

	Df	Deviance	Resid. Df	Resid. Dev.	P value
NULL			6399	16761	
fStep	9	676.13	6398	16085	< 0.001
vStep	9	1830.78	6397	14254	< 0.001
Group	2	72.3	6395	14182	< 0.001
fStep:vStep	19	0	6394	14182	0.95
fStep:Group	2	330.4	6392	13851	< 0.001
vStep:Group	2	659.9	6390	13191	< 0.001
fStep:vStep:Group	2	8.67	6388	13183	< 0.05

Table 1. Analysis of deviance table from model.

Further models were fitted within each group explore the interactions with Group (Tables 3a-c). The results show a significant effect for the vocalic dimension for all groups, but an effect for frication dimension in only the frication and the pinyin groups. Figure 1 give a graphic representation of the data.

Vocalic	Df	Deviance	Resid. Df	Resid. Dev.	P value
NULL			2099	6124.5	
fStep	9	0.52	2098	6124	0.47
vStep	9	2028.46	2097	4095.5	< 0.001

Table 2a. Vocalic group.

Frication	Df	Deviance	Resid. Df	Resid. Dev.	P value
NULL			1799	3544	
fStep	9	491.38	1798	3052.6	<2e-16
vStep	9	89.92	1797	2962.7	<2e-16

Table 2b. Frication group.

Pinyin	Df	Deviance	Resid. Df	Resid. Dev.	P value
NULL			2499	7025.7	
fStep	9	497.23	2498	6528.5	< 0.001
vStep	9	395.15	2497	6133.4	< 0.001

Table 2c. Pinyin group.

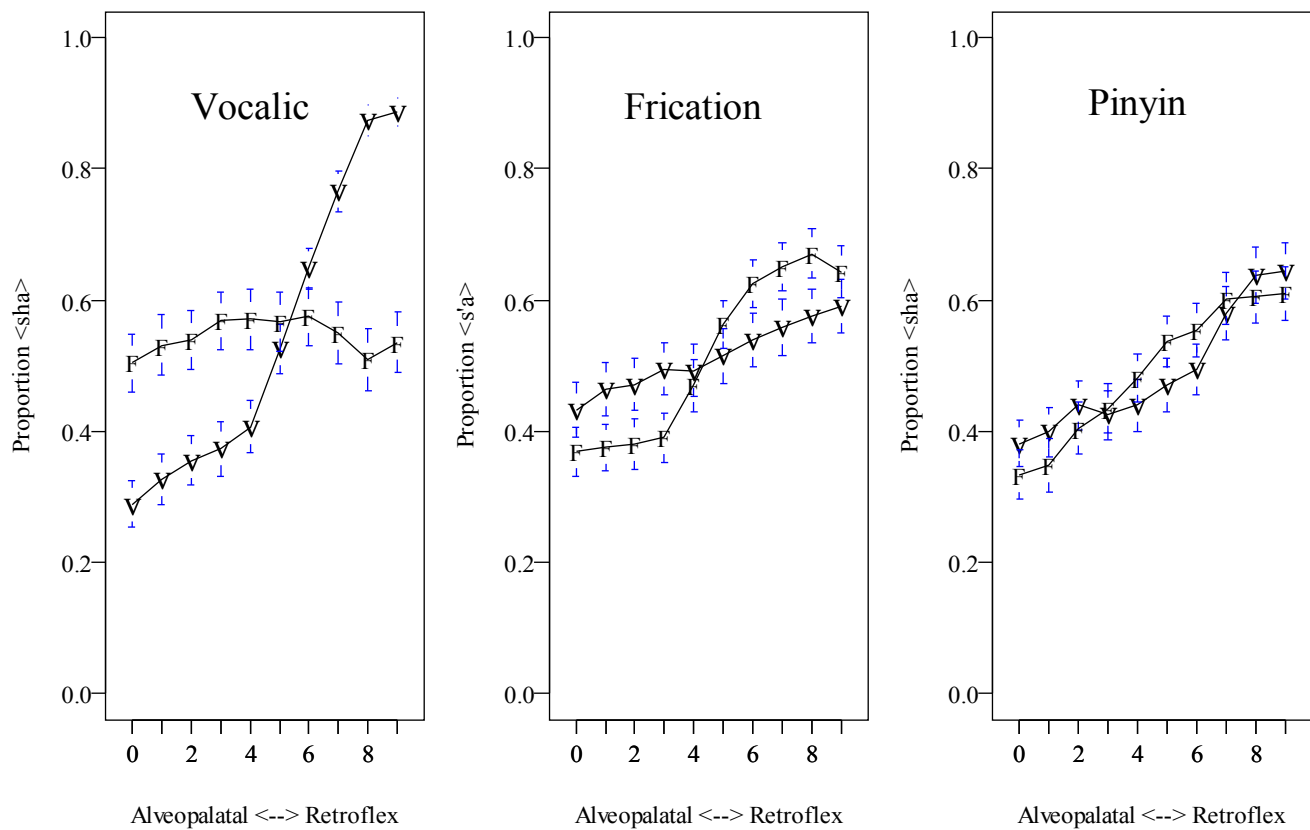


Fig. 1. Proportion $\langle sha \rangle / \langle s'a \rangle$ responses for each group. F = frication dimension, V = vocalic dimension. Leftmost panel represents the vocalic group, middle panel the frication group, rightmost panel the pinyin group.

Both the vocalic dimension in the vocalic group and the frication dimension in the frication group show effects that resemble classic categorical perception (i.e. “s” curves). The frication curve is weak, largely due to high variance among the subjects. In the vocalic group, there is general tendency to respond “sha”. In the pinyin group, there results are primarily linear, suggesting categorization, but not categorical perception. Individual subject plots demonstrate that there was a high degree of variance in the use dimensions by this group such that some subjects seem to rely on frication noise, others on the vocalic information, and some used both. None were especially consistent, but as a group they combined to the result above.

3. Discussion

Overall, the expected effects are present; the orthography had a clear effect with very little exposure and no explicit instructions on what cues to attend to. In all groups the vocalic dimension was given some weight, though only the vocalic group showed a classic categorical perception curve. This suggests that there is strong preference to weight the vocalic dimension and that the orthographic labels further reinforced this weighting. A secondary orthographic effect is that the pinyin and vocalic groups both had a symbol for the English post alveolar sibilant $\langle sh \rangle$. In the case of the vocalic group this may have led to a

general bias to use that label and no subjects using that label for the alveolopalatal fricative (i.e. no need for adjustment).

Overall, the effects were not incredibly dramatic, though they were reliable. This probably reflects the extremely limited exposure; listeners only had 45 minutes with the sounds and only 5 opportunities to classify any one stimulus. The pinyin group showed the most variability and may have had the most difficult task. This may be due to having to spread attentional resources across the syllable rather than concentrating them on one or the other cue.

4. Conclusions

This study demonstrated orthographic effects on cues, below the segment level. Listener's cue weighting strategies were changed as a result of the label set they were asked to use. This ties into the general view that the orthographic system learned by the reader drives how much internal structure speakers derive from the speech signal. While this may argue against the phoneme as a universal concept (Twaddell 1935, Firth 1948, Faber 1992, Öhman 2000, Port and Leary 2005), at least from a metalinguistic standpoint, it does not deny that such structure is available to listeners when directed to it. Perhaps more importantly such results instead argue that the extremely discrete, hierarchical view of phonetic structure should not be the default assumption when describing and examining perceptual targets (and presumably articulatory ones, too), regardless of the utility that such conceits as the International Phonetic Alphabet provide.

Relatedly, the results of this study call into question the causes of the weighting shift seen in English learning children acquiring its sibilant contrast. That the shift happens in the early stages of learning to read (6-8yrs), that weighting shifts from vocalic to frication (consonantal) cues, that it is related to PA, and that orthography affects PA and weighting all combine to argue that this weighting shift happens as a result of learning to read. Indeed, the research on fricative cue use, especially that which demonstrates listeners' reliance on fricative noise (e.g. Wagner et al. 2006) is focused solely on western languages having alphabetic systems. The primary ramification of this, if true, is that cue weighting may not be solely determined by salience and reliability, but that other factors may conspire to change weighting strategies.

Ultimately, both of these points argue that a more complete understanding of language use must be derived from languages and populations different from the traditional literate western alphabetical tradition. While certainly not a new concept (e.g. Sears 1986), the results and views outlined in this paper add to the view that linguistic research must be made more diverse.

5. References

Bentin, S. (1992). Phonological awareness, reading, and reading acquisition: A survey and appraisal of

- current knowledge, in R. Frost & L. Katz (eds.), *Orthography, Phonology, Morphology, and Meaning*, Elsevier Science Publishers B.V., Amsterdam, pp. 193-210.
- Bentin S, Hammer R, Cahan S. (1991). The effects of aging and first-grade schooling on the development of phonological awareness. *Psychological Science*, 2, 271–274.
- Bertelson, P., de Gelder, B., Tfouni, L.V., & Morais, J. (1989). Metaphonological abilities of adult illiterates: New evidence of heterogeneity. *European Journal of Cognitive Psychology*, 1(3), 239–250.
- Bruce, D.J. (1965). The analysis of word sounds. *British Journal of Educational Psychology*, 34, 158-170.
- Cheung, H., Chen, H.-C., Lai, C.Y., Wong, O.C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition*, 81, 227–241.
- Cheung, H., & Chen, H.-C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes*, 19, 1-28.
- Faber, A. (1992). Phonemic segmentation as epiphenomenon: Evidence from the history of alphabetic writing. In P. Downing, S. Lima & M. Noonan (Eds.), *The Linguistics of Literacy* (pp. 111-134). Amsterdam: John Benjamins.
- Firth, J. R. (1948). Sounds and prosodies. *Transactions of the Philological Society*, 127-152.
- Johnson, K. (2003). *Acoustic and Auditory Phonetics* (2nd ed.). Malden, MA: Blackwell.
- Kirtley, C., Bryant, P., MacLean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology*, 78, 243-255.
- Lyon, R. (1998). Overview of Reading and Literacy Initiatives: Report to Committee on Labor and Human Resources, United States Congress.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- McGuire, G. (2007). Phonetic category learning. Unpublished dissertation. The Ohio State University.
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7, 323-331.
- Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, 24, 45–64.
- Nittrouer, S. (1992). Age-related differences in perceptual effects for formant transitions within syllable and across syllables. *Journal of Phonetics*, 20, 1-32.
- Nittrouer, S. (1996). The relation between speech perception and phonemic awareness: Evidence from low SES children and children with chronic OM. *Journal of Speech and Hearing Research*. 39, 1059-1070.
- Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America*. 112(2), 711-719.

- Nittrouer, S. & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *The Journal of the Acoustical Society of America*, 101(4), 2253-2266.
- Nittrouer, S., and Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, 30, 319-329.
- Öhman, S. (2000). Expression and content in linguistic theory. In M. Gustafsson & L. Hertzberg (Eds.), *The Practice of Language*. Dordrecht: Kluwer.
- Port, R., & Leary, A. (2005). Against symbolic phonology. *Language*, XXX.
- Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, 24, 31-44.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3), 515-530.
- Twaddell, W. F. (1935). On defining the phoneme. *Language*, Language Monograph 16.
- Wagner, A., Ernestus, M., & Cutler, A. (2006). Formant transitions in fricative identification: The role of native fricative inventory. *The Journal of the Acoustical Society of America*, 120(4), 2267-2277.