



MONTCLAIR STATE
UNIVERSITY

Montclair State University
**Montclair State University Digital
Commons**

Department of Psychology Faculty Scholarship
and Creative Works

Department of Psychology

7-2-2016

Catching the Drift: Carol A. Fowler on Phonetic Variation and Imitation

Jennifer Pardo

Montclair State University, pardoj@montclair.edu

Follow this and additional works at: <https://digitalcommons.montclair.edu/psychology-facpubs>



Part of the [Psychology Commons](#)

MSU Digital Commons Citation

Pardo, Jennifer, "Catching the Drift: Carol A. Fowler on Phonetic Variation and Imitation" (2016).

Department of Psychology Faculty Scholarship and Creative Works. 119.

<https://digitalcommons.montclair.edu/psychology-facpubs/119>

This Article is brought to you for free and open access by the Department of Psychology at Montclair State University Digital Commons. It has been accepted for inclusion in Department of Psychology Faculty Scholarship and Creative Works by an authorized administrator of Montclair State University Digital Commons. For more information, please contact digitalcommons@montclair.edu.

RUNNING HEAD: Catching the Drift

Catching the Drift: Carol Fowler on Phonetic Variation and Imitation

Jennifer S. Pardo
Montclair State University
Psychology Department

Correspondence:
Jennifer S. Pardo
Montclair State University
Psychology Department
1 Normal Avenue
Montclair, NJ 07043
pardoj@montclair.edu

Abstract

In Carol Fowler's Direct Realist account of speech perception, linguistically significant gestures of the vocal tract are a common currency for both speech perception and production. A straightforward prediction of this account is that listeners will produce what they perceive, leading to imitation or gestural drift. Many studies by Fowler and colleagues have established gestural imitation across acoustic, perceptual, and articulatory measures, and provided a valuable framework for understanding phonetic form variation and imitation. As such, this framework's enduring legacy will continue to enrich an understanding of phonetic form variation in spoken communication. This paper reviews Fowler's pioneering work on these issues and some of the work that it has inspired.

Catching the Drift: Carol Fowler on Phonetic Variation and Imitation

In Carol Fowler’s Direct Realist account of speech perception, linguistically significant gestures of the vocal tract are a common currency in speech production and perception. Vocal tract gestures satisfy at least two important constraints in linguistic communication—they are public physical events, and their activity can distinguish words in a language (Fowler, 2010, 2014; Goldstein & Fowler, 2003). From a Direct Realist perspective, vocal tract gestures cause patterning in informational media and perception directly resolves gestures from patterned physical energy, whether acoustic, visual, or even haptic (Fowler, 1986). To the extent that perception directly resolves vocal tract gestures, their properties would be available to a listener when subsequently producing speech. This circumstance leads to a straightforward proposal that talkers might imitate the phonetic forms they hear, ultimately contributing to phonetic drift over time.

Fowler’s first study of phonetic drift was prompted by anecdotal reports of long-term changes in native phonetic repertoires of adults with exposure to either a different dialect or a different language environment (Sancier & Fowler, 1997). After living in America for long periods of time, a British English speaker adopted a “ghastly” American accent according to his compatriots in England, and a Brazilian Portuguese speaker likewise elicited critical comments from her father about her “explosive” speech. Sancier and Fowler (1997) focused on the latter case, that of a second language environment influencing pronunciation of consonants in a talker’s first language production. In this case, the explosiveness of the talker’s speech was hypothesized to arise due to differences in voice onset timing (VOT) in English (L2) and Brazilian Portuguese (L1)—after experiencing American English consonants during long stays in the US, her production of

consonants in both languages had drifted in pronunciation toward the norms of American English.

Voice onset timing is a particularly useful attribute for this endeavor because it results from a single vocal tract gesture that has consequences for lexical identification, and its acoustic consequences are relatively easy to measure and manipulate. For example, the words *bit* and *pit* in English differ only in relative timing of a devoicing gesture, with longer VOTs for *pit* than for *bit*. Because VOTs of voiceless stops in Brazilian Portuguese are typically shorter than VOTs of corresponding consonants in English, producing Brazilian Portuguese words with longer VOTs could contribute to a so-called explosive impression in a native listener. Sancier and Fowler (1997) first confirmed that naïve native Brazilian Portuguese listeners judged the talker's Portuguese as more accented after a 4-month stay in the US than after a 2.5-month stay in Brazil. Acoustic measures of words beginning with /t/ and /p/ in both languages revealed that VOTs drifted toward the norms of English after long stays in the US (lengthening by 5 to 6 ms on average) relative to their values after a stay in Brazil. Other researchers have found similar patterns when examining long-term changes in vowel formants, that is, talkers show long-term adaptation to different dialects and even college roommates (Evans & Iverson, 2007; Pardo, Gibbons, Suppes, & Krauss, 2012).

This study sparked an ongoing line of research on phonetic imitation of VOT that has led to at least four major discoveries (Abrego-Collier, Grove, Sonderegger, & Yu, 2011; Chang, 2012, 2013; Fowler, Brown, Sabadini, & Weihing, 2003; Nielsen, 2011; Olmstead, Viswanathan, Aivar, & Manuel, 2013; Sanchez, Miller, & Rosenblum, 2010; Shockley, Sabadini, & Fowler, 2004; Yu, Abrego-Collier, & Sonderegger, 2013). First,

talkers adapt rapidly to model utterances in a speech shadowing task, producing forms that converge in VOT to auditory prompts at latencies that approximate simple reaction times (Fowler et al., 2003). Moreover, spontaneous imitation of VOTs of heard forms can be detected by naïve listeners who judge shadowed items to be more similar to auditory prompts from model talkers than baseline items in an AXB similarity task (Shockley et al., 2004). Furthermore, audiovisual presentation likewise elicits VOT convergence, indicating that gestural information is preserved visually as well as acoustically (for /p/: Sanchez et al., 2010). Finally, personality attributes and a talker's attitude toward a model talker influence degree of VOT imitation (Yu et al., 2013).

A follow-up study with a much larger set of talkers examined cross-language effects on VOTs in English and French among monolinguals and bilinguals in Montreal (Fowler, Sramko, Ostry, Rowland, & Hallé, 2008). Like Brazilian Portuguese, VOTs of voiceless stops in French are shorter than corresponding phonemes in English. Although the study did not assess phonetic drift or imitation directly, Fowler et al. found that VOTs of French /p/, /t/, and /k/ produced by simultaneous French-English bilinguals were longer than those of monolingual French speakers. Likewise, VOTs of English /p/, /t/, and /k/ produced by simultaneous French-English bilinguals were shorter than those of monolingual English speakers. These patterns favor an account of cross-language phonetic influences, possibly via sharing of a devoicing vocal tract gesture.

Importantly, Fowler has extended this work to other aspects of behavior. Honorof, Weihing, and Fowler (2011) found imitation of l-variants in both acoustic and articulatory measures. That is, talkers who rapidly shadowed VCVs containing either light or dark versions of /l/ produced different forms that approximated the model's

versions. Nye and Fowler (2003) found that shadowers imitated pseudoword patterns more frequently when their phonotactic properties conformed less to English, perhaps because the unfamiliar items induced greater attention to details of the utterances, which were then imitated. In an intriguing set of experiments, work with Shockley and colleagues found that interacting talkers entrained their postural sway during conversation, exhibiting greater shared positions between talkers who spoke to each other versus to another person (Shockley, Santana, & Fowler, 2003). Moreover, follow-up experiments established that interpersonal coordination of postural sway patterns results specifically from sharing of articulatory processes (Baker, Shockley, Richardson, & Fowler, 2005; Fowler, Richardson, Marsh, & Shockley, 2008; Shockley, Baker, Richardson, & Fowler, 2007).

Phonetic imitation and drift is well accounted for by proposing shared forms in speech production and perception. Other accounts likewise propose sharing of representations and/or processes that could support imitation (see Liberman, 1996; Pickering & Garrod, 2013). The notion of vocal tract gestures as common currency in speech perception and speech production has been supported by findings that areas of brain activation during speech perception and production overlap, and in particular, activation of the inferior frontal gyrus was greater when talkers tried to imitate speech than when perceiving alone (Irwin, Frost, Mencl, Chen, & Fowler, 2011). However, Fowler has long acknowledged limitations to phonetic imitation. In Sancier and Fowler (1997), VOTs in Brazilian Portuguese drifted (lengthened), but did not match the distributions of VOTs in English (and vice versa for English VOTs). At the time, they proposed a dynamical systems account of phonological categories in which phonemes are

characterized by potential function attractors of differing strengths. A potential function attractor specifies phase relations among vocal tract gestures involved in a particular phoneme, such as relative timing of devoicing and bilabial closure gestures for /p/. In this talker's case, her primary Brazilian Portuguese attractor for /p/ would exert a stronger influence on production than later-acquired attractors (such as for English /p/), resulting in drift rather than perfect imitation.

Failures to find perfect matching/imitation of phonetic forms are the norm in the literature (see review in Pardo, Jordan, Mallari, Scanlon, & Lewandowski, 2013). Imperfect imitation arises from multiple possibly competing tendencies—a propensity to imitate, persistence in habitual speaking patterns, a talker's situational goals, and social aspects of conversational settings (Fowler, 1986, 2010, 2014; Pardo, 2006). Although spoken communication relies on the establishment of parity, shared forms need not be completely shared: “That gestures are only *sufficiently* shared, not wholly shared, among speakers of a common language underlies language change and serves a valuable function of expressing social identity” (Fowler, 2014, p. 177). Accordingly, phonetic forms serve multiple roles of specifying linguistic tokens, contributing to interpersonal coordination, and expressing social identity.

The legacy of Carol Fowler's approach is secure. She has provided an elaborate yet parsimonious account of speech perception as it relates to speech production, situating the account in a broad framework. Her Direct Realist perspective has always acknowledged multiple affordances of speech forms, for specifying linguistic aims and for coordinating individuals, among other functions. As such, this framework's enduring legacy will continue to enrich an understanding of phonetic form variation in spoken

communication. By turning to notions of entrainment of dynamical systems, this approach points to promising avenues for future research.

Author Note

Partial support for completion of this paper was provided by NSF PAC grant #BCS-1229033 to Jennifer Pardo (PI). The author thanks Jim Magnuson, Robert Remez, and Navin Viswanathan for their helpful comments and support.

References

- Abrego-Collier, C., Grove, J., Sonderegger, M., & Yu, A. C. L. (2011). Effects of speaker evaluation on phonetic convergence. *Proceedings of the International Congress of the Phonetic Sciences XVII*.
- Baker, A. A., Shockley, K. D., Richardson, M. J., & Fowler, C. A. (2005). Verbal constraints on interpersonal postural coordination. In H. Heft & K. L. Marsh (Eds.), *Studies in perception and action VIII*. (pp. 41-44). Mahwah, NJ: Lawrence Erlbaum Associates.
- Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*, 249-268.
doi:10.1016/j.wocn.2011.10.007
- Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, *41*, 520-533. doi:10.1016/j.wocn.2013.09.006
- Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: a study of accent change in young adults. *Journal of the Acoustical Society of America*, *121*, 3814-3826. doi:10.1121/1.2722209
- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, *14*, 3-28.
- Fowler, C. A. (2010). Embodied, embedded language use. *Ecological Psychology*, *22*, 286-303. doi:10.1080/10407413.2010.517115

- Fowler, C. A. (2014). Talking as doing: Language forms and public language. *New Ideas in Psychology*, 32, 174-182. doi:10.1016/j.newideapsych.2013.03.007
- Fowler, C. A., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, 49(3), 396-413. doi:10.1016/S0749-596X(03)00072-X
- Fowler, C. A., Richardson, M. J., Marsh, K. L., & Shockley, K. D. (2008). Language use, coordination, and the emergence of cooperative action. In A. Fuchs & V. K. Jirsa (Eds.), *Coordination: Neural, Behavioral and Social Dynamics* (pp. 261-279). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-74479-5
- Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. A. (2008). Cross language influences on the speech of French-English bilinguals. *Journal of Phonetics*, 36, 649-663.
- Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In A. S. Meyer & N. O. Schiller (Eds.), *Phonetics & Phonology in Language Comprehension & Production: Differences & Similarities* (pp. 1-53). Berlin: Mouton.
- Honorof, D. N., Weihing, J., & Fowler, C. A. (2011). Articulatory events are imitated under rapid shadowing. *Journal of Phonetics*, 39(1), 18-38.
doi:10.1016/j.wocn.2010.10.007
- Irwin, J. R., Frost, S. J., Mencl, W. E., Chen, H., & Fowler, C. A. (2011). Functional activation for imitation of seen and heard speech. *Journal of Neurolinguistics*, 24,

611-618. doi:10.1016/j.jneuroling.2011.05.001

Liberman, A. M. (1996). *Speech: A Special Code*. Cambridge, MA: MIT Press.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39(2), 132-142. doi:10.1016/j.wocn.2010.12.007

Nye, P. W., & Fowler, C. A. (2003). Shadowing latency and imitation: The effect of familiarity with the phonetic patterning of English. *Journal of Phonetics*, 31, 63-79.

Olmstead, A. J., Viswanathan, N., Aivar, M. P., & Manuel, S. (2013). Comparison of native and non-native phone imitation by English and Spanish speakers. *Frontiers in Psychology*, 4, 475. doi:10.3389/fpsyg.2013.00475

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119(4), 2382-2393. doi:10.1121/1.2178720

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40(1), 190-197.
doi:10.1016/j.wocn.2011.10.001

Pardo, J. S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013). Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory & Language*, 69, 183-195.
doi:10.1016/j.jml.2013.06.002

Sanchez, K., Miller, R. M., & Rosenblum, L. D. (2010). Visual influences on alignment to voice onset time. *Journal of Speech, Language, and Hearing Research*, 53(2), 262-

272.

- Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, *25*, 421-436.
- Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception & Performance*, *33*, 201-208. doi:10.1037/0096-1523.33.1.201
- Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Attention, Perception, & Psychophysics*, *66*, 422-429.
- Shockley, K., Santana, M., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 326-332. doi:10.1037/0096-1523.29.2.326
- Yu, A. C. L., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “Autistic” traits. *PLoS ONE*, *8*, e74746. doi:10.1371/journal.pone.0074746