Journal of Phonetics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

The representation and execution of articulatory timing in first and second language acquisition

Melissa A. Redford [a,*], Grace E. Oh [b]

[a] University of Oregon, United States
[b] Konkuk University, South Korea

ARTICLE INFO

Article history:
Received 15 December 2015
Received in revised form 3 January 2017
Accepted 12 January 2017
Available online xxxx

Keywords:
Vowel duration
Temporal variability
Temporal patterns
Speech motor control
Speech acquisition

ABSTRACT

The early acquisition of language-specific temporal patterns relative to the late development of speech motor control suggests a dissociation between the representation and execution of articulatory timing. The current study tested for such a dissociation in first and second language acquisition. American English-speaking children (5- and 8-year-olds) and Korean-speaking adult learners of English repeatedly produced real English words in a simple carrier sentence. The words were designed to elicit different language-specific vowel length contrasts. Measures of absolute duration and variability in single vowel productions were extracted to evaluate the realization of contrasts (representation) and to index speech motor abilities (execution). Results were mostly consistent with a dissociation. Native English-speaking children produced the same language-specific temporal patterns as native English-speaking adults, but their productions were more variable than the adults'. In contrast, Korean-speaking adult learners of English typically produced different temporal patterns than native English-speaking adults, but their productions were as stable as the native speakers'. Implications of the results are discussed with reference to different models of speech production.

1. Introduction

Articulatory timing refers to the coordination of speech articulators in time to achieve motor goals in sequence. Given this definition, timing can be thought of either as a motor speech skill or as a language behavior: stable coordination patterns emerge with neuromotor maturation and speech motor practice; goal sequencing emerges with the acquisition of language. Whereas children are slow to acquire stable coordination patterns (Smith & Zelaznik, 2004), language-specific sequencing is acquired fairly early (Stoel-Gammon & Dunn, 1985). These observations suggest a dissociation between the representation and execution of timing information, consistent with a theoretical distinction between competence (representation) and performance (execution). The current study tested this dissociation hypothesis against an alternative interaction hypothesis by investigating the effect of language-specific vowel length contrasts on production in first and second language acquisition.

1.1. Development of timing control

Skilled action includes patterns of movement coordination that are acquired for functional ends (i.e., goals). Change in the duration and variability with which goal-directed movements are executed is thought to reflect neuromotor maturation and/or motor learning (see Smith, 1992). Whatever the underlying explanation, both duration and variability are observed to decrease as coordinated articulatory movements become faster and more stable (Smith & Zelaznik, 2004).

The earliest studies to link acoustic duration and temporal variability in children's speech to motor skill development focused on linguistic units of various sizes, including segments, syllables, and words (Kent & Forner, 1980; Smith, 1978; Tingley & Allen, 1975). For example, Smith (1978) compared the mean acoustic duration and standard deviation of repeated word productions in 2- and 4-year-old children's speech to adults' speech. He found that children's word durations were greater than adults', and that 2-year-olds' repetitions of the same word were more variable than adults'. Kent and Forner (1980) found that even 6-year-olds produced more variable phrase, word, and segment durations than adults. Noting the correlation between mean duration

* Corresponding author at: Linguistics Department, 1290 University of Oregon, Eugene, OR 97403, United States.
E-mail address: redford@uoregon.edu (M.A. Redford).

http://dx.doi.org/10.1016/j.wocn.2017.01.004
0095-4470/

and standard deviation in their data, Kent and Forner examined whether age-related differences would persist if standard deviations were mean normalized. They did, leading the authors to conclude that duration and temporal variability were independent markers of motor skill in children's speech. This conclusion has since been echoed many times in developmental studies of speech production (e.g., Lee, Potamianos, & Narayanan, 1999; Redford, 2014; Smith, 1992; Smith, Sugarman, & Long, 1983), and is consistent with the broader literature on motor learning (see, e.g., Rosenbaum, 2009: Ch. 4).

Developmental studies have also found that children's speech is more acoustically variable than adult speech until age 12 years (e.g., Lee et al., 1999); kinematic differences persist until age 14 years (e.g., Green, Moore, Higashikawa, & Steeve, 2000; Sharkey & Folkins, 1985; Smith & Goffman, 1998; Smith & Zelaznik, 2004). These findings indicate that speech motor development is protracted. Despite this, children produce linguistically-relevant temporal patterns accurately from a very early age. For example, in the aforementioned study on speech timing in 2-year-old children, Smith (1978) also investigated the effects of place of articulation and voicing on the children's production of mean stop closure duration, VOT duration, and vowel duration. He found that whereas absolute duration values differed in child and adult speech, the proportional duration of these intervals varied with linguistic factors in the same way across all age groups. He concluded from these and other results from the same study that "even prior to age three, children recognize important temporal parameters of the language they are learning and incorporate them into their phonological system—a system which, despite certain limitations, seems quite sophisticated (p. 65)."

Subsequent acoustic–phonetic studies on early child language have confirmed the idea that children acquire temporal information early as part of their language grammar or abstract word form representations. For example, a number of studies on lexical stress production in young children have shown that children use duration to distinguish stressed from unstressed syllables in English as early as 2 years of age (e.g., Kehoe, Stoel-Gammon, & Buder, 1995; Pollock, Brammer, & Hageman, 1993; Schwartz, Petinou, Goffman, Lazowski, & Cartusciello, 1996). Studies on stop production indicate that English-speaking children use voice onset time to convey a voicing contrast from an early age (Bond & Wilson, 1980; Imbrie, 2005), even if during the earliest period (prior to age 2 years) the contrast is not perceptible to adults (Macken & Barton, 1980). Two year old children also use vowel duration to reliably signal voicing in stop codas (Buder & Stoel-Gammon, 2002; Song, Demuth, & Shattuck-Hufnagel, 2012).

In sum, studies on speech motor development have shown that children's speech is slower and more variable than adults' speech, and that this difference persists until at least age 12. In contrast, studies on early child phonology indicate that language-specific temporal patterns are mostly acquired by 3 years of age. Some difficulties in specific sound or cluster production persist until children have begun school at age 5 (see, e.g., Stoel-Gammon & Dunn, 1985), but resolve soon thereafter and certainly well before speech motor abilities are adult-like.

1.2. Timing control in second language acquisition

Whereas children acquire language-specific timing patterns early during first language (L1) acquisition, adult second language (L2) learners often fail to achieve native-like timing in their L2; instead, they produce L1 influenced patterns. For example, adult Spanish-speaking learners of English produce voiceless stops in English with shorter voice onset times (VOT) than do native English speakers (Flege, 1991), presumably because Spanish voiceless stops are characterized by shorter VOTs than English voiceless stops in syllable positions where these are released. Adult Korean-speaking learners of English produce less contrastive vowel durations to signal differences in coda stop voicing than native English speakers (Cho & Shin, 2013), presumably because the Korean voicing contrast for stops is neutralized in final position. Zsiga (2003) cites many similar examples and goes on to show that specific patterns of Russian word-to-word timing influence Russian learner's production of English patterns. She also reports that English learners' of Russian produce unmarked articulatory timing patterns that do not occur in either English or Russian. Zsiga interprets the former results to support the notion of cross-language transfer and the latter to support the idea of distinct second language representations that may also reference universal phonological processes (cf. Selinker, 1972).

It is the interpretation of the L2 findings that is especially relevant to our present interest in a dissociation between the representation and execution of articulatory timing. Non-native timing patterns in second language speech are nearly always explained with reference to representational factors; not motoric ones. This is true even when the observed patterns cannot be explained either in terms of the L1 or L2 patterns, as in the Russian learner results reported in Zsiga (2003; see also Cebrian, 2000). Moreover, Flege (1991:406) explicitly rejects the idea that adult learners are less able than early learners "to motorically implement their perceptual representations for sounds," noting with reference to his VOT data that there "is no *a priori* reason to think that it is somehow easier for late learners to produce partial modification of previously established articulatory patterns than to produce a *complete* modification that would enable them to match native speakers of English (emphasis in the original)." The implication is that motor factors have no impact on L2 representations, in keeping with the dissociation hypothesis. Relatedly, the two main theories of second language speech acquisition, the Speech Learning Model (SLM; Flege, 1995) and Perceptual Assimilation Model (PAM; Best, 1995), are models of perceptual learning; specifically, they are models of how pre-existing phonological categories influence and are influenced by second language speech perception. Motor learning and control are not considered in the models.

In sum, the mainstream assumption in adult second language acquisition research is that the motor system faithfully executes whatever "interlanguage" representation has been established. This is likely both because foreign accents tend to be stable over time and because we often think of speech motor skills in maturational terms.

Please cite this article in press as: Redford, M. A., & Oh, G. E. The representation and execution of articulatory timing in first and second language acquisition. *Journal of Phonetics* (2017), http://dx.doi.org/10.1016/j.wocn.2017.01.004