Journal of Phonetics 71 (2018) 98–112

Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

The acquisition of syllable-level timing contrasts by English- and Spanish-speaking bilingual children with normal hearing and English- and Spanish-speaking bilingual children with cochlear implants

Mark Gibson [a,*], Elizabeth Goodin-Mayeda [b], Ferenc Bunta [b], Arturo Hernandez [b]

[a] *Universidad de Navarra, Spain*
[b] *University of Houston, United States*

A R T I C L E I N F O

Article history:
Received 8 January 2018
Received in revised form 14 July 2018
Accepted 20 July 2018

Keywords:
Syllable timing
Spanish
Bilingual acquisition
Phonology
Cochlear implant users

A B S T R A C T

We examined the Spanish productions of four syllable-level timing parameters by bilingual English- and Spanish-speaking children with normal hearing (NH group, henceforth) and bilingual English- and Spanish-speaking children with cochlear implants (CI group, henceforth) during the stages of phonological development. The four temporal variables related to syllable structure include: (1) lateral duration in simplex and complex onsets, (2) vowel duration following simplex and complex onsets, (3) voice onset time (VOT, henceforth) for stops in simplex and complex onsets, and (4) interplateau intervals (IPI, henceforth), or the latency between the release of articulatory constriction of the first consonant (C1, henceforth) in a C1C2 sequence and the onset of articulatory constriction of the second consonant (C2, henceforth). Our motivation for the study is to address (1) how, and at what age, children acquire and produce the temporal patterns related to syllable structure of their native language(s), (2) the effects of language interference in the acquisition of timing parameters by bilingual children, and (3) whether hearing impairment impedes correct acquisition and categorization of temporal cues related to syllables. Results of the Spanish production tasks show that the NH groups' productions were in general commensurate with data for monolingual productions, and the results of the CI group's productions show general concurrence with the NH group. However, some differences with monolingual timing patterns, and effects based on group, were found which will be discussed in the context of existing studies dealing with the acquisition of timing parameters related to syllables.

1. Introduction

Previous research in speech production has revealed a number of effects of high-level syllable structure on articulatory timing. The acoustic products of these articulatory patterns have important implications for phonological development, and may be useful to children in constructing early perceptual templates (as per Gobet & Simon, 2000) regarding syllable structure. It is generally accepted that during the stages of phonological acquisition, children access the phonetic and phonological categories of their native language by way of acoustic-level cues they perceive from the statistical distribution of the sounds of their native language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Maye, Werker, & Gerken, 2002). Few studies, however, have investigated how subjects with a distorted audio input, say children with cochlear implants, acquire the syllable-level timing parameters that typify their native language(s), and how this distortion of the audio signal affects early productions.[1] There is good evidence to suggest that unidimensional temporal cues related to syllable-level timing such as VOT are perceived and acquired well by children with cochlear implants (Bunta, Goodin-Mayeda, Procter, & Hernandez, 2016; Caldwell & Nittrouer, 2013;

* Corresponding author at: 1330 Edificio de Bibliotecas, Universidad de Navarra, Campus Universitario, s/n, 31009 Pamplona, Spain.
E-mail address: mgibson@unav.es (M. Gibson).

[1] Our working assumption is that the CI children's productions are in some cases causally related to their ability to accurately perceive cues from the phonetic environment. We use this assumption as a starting ground to address our topic, but we do not mean to suggest that accurate perception is the only important factor in the CI groups' phonological development. There is a dearth of evidence, for example, to suggest that visual information is also used to perceive some aspects of speech. Also, in some instances, age of implantation can explain many aspects of language development in CI users better than just the limitations or effectiveness of the hearing device.

https://doi.org/10.1016/j.wocn.2018.07.005
0095-4470/

Donaldson, Rogers, Johnson, & Oh, 2015; Iverson, Smith, & Evans, 2006), yet relatively little is known regarding the acquisition of the temporal properties of multiple acoustical dimensions that are interconnected within the broader unit of the syllable. Concurrently, scant attention has been given in the literature to how differing temporal parameters between languages affect early phonological development and production in bilingual populations. The present study aims to address these gaps in the research by examining the Spanish productions (English productions will be taken up in a future study) of syllable-level timing patterns by bilingual English and Spanish speaking children with cochlear implants and bilingual English and Spanish speaking children with normal hearing.

A number of studies have shown that certain temporal characteristics of consonants and vowels vary according to their position in a syllable. Consider, for example, the behavior of laterals in different onset types. In a number of languages such as Spanish and English, lateral duration has been found to be shorter when appearing in complex onsets, as compared to word initial singleton lateral onsets (Gibson, Sotiroupoulou, Tobin, & Gafos, 2017; Haggard, 1973; Hermes, Grice, Mücke, & Niemann, 2008; Marin, 2013; O'Shaughnessy, 1974). Thus, lateral duration reduces as syllable complexity increases (as more consonants are added to the onset). In a similar way, the compensatory shortening, or compression, of the syllable nucleus has also been found to be a function of onset complexity (Hermes et al., 2008; Marin & Bučar Shigemori, 2014; Marin & Pouplier, 2010; Munhall, Fowler, Hawkins, & Saltzman, 1992), whereby vowels appearing after complex onsets are shorter than those appearing after singleton onsets.

The syllable position of stops has also been shown to have effects on oral-laryngeal timing, which may have important implications for phonological development across different groups. In English, for example, word initial voiceless stops preceding stressed vowels are characterized by a long voice onset time (the latency from the release of articulatory constriction of the stop to the onset of phonation, VOT henceforth) vis-à-vis word internal stops preceding unstressed vowels (Lisker & Abramson, 1964). Hence the initial /p/ in the word 'paper' exhibits a longer VOT than the word internal /p/. In clusters, VOT for the voiceless stop (for example in the word 'spark') is also reduced, patterning more like VOT for a voiceless [b̥] in English. Relatedly, voiced stops in phrase initial position in English and other Germanic languages often undergo devoicing, such that a voiceless [b̥], for example, is only distinguishable from a voiceless [p] by way of VOT.

With regard to cross-linguistic patterns of VOT, two typological patterns have been reported for bimodal voicing contrasts (Maddieson, 2009). Languages such as French and Spanish contrast phonologically voiced and voiceless stops by way of a true voice versus short lag VOT paradigm, such that phonologically voiced stops are typically realized as (pre)voiced, while voiceless stops are typified by a short lag VOT (Caramazza & Yeni-Komshian, 1974; Cho and Ladefoged, 2000; Lisker & Abramson, 1964; Rosner, López-Bascuas, & García-Albea, 2000; Williams, 1977). Alternatively, languages such as German and English maintain the voice contrast for initial singleton stops (in stressed syllables) by way of a short lag versus long lag VOT pattern, whereby voiced stops are characterized by a short lag VOT and voiceless stops are typified by a longer VOT in relation to voiced stops (Bombien & Hoole, 2013; Hoole, Bombien, Kühnert, & Mooshammer, 2009; Klatt, 1975). Thus, while in Spanish and French, the phonological voicing contrast in stops is expressed as a function of voicing (either (pre)voiced or voiceless), other languages such as English and German (Klatt, 1975), where devoicing of the voiced stop may occur in some contexts, the phonological voicing contrast is expressed phonetically by way of VOT.

Place of articulation has long been reported to affect VOT, velar consonants having longer VOTs than labials and coronals (Crystal & House, 1988; Docherty, 1992; Hoole, 2006; Lisker & Abramson, 1964; Maddieson, 1997; Marin & Bučar Shigemori, 2014; Nearey & Rochet, 1994; Weismer, 1980). The biomechanical basis underlying the effects of place of articulation on VOT has been extensively discussed in Maddieson (1997), Cho and Ladefoged (2000), Stevens (2000), and Bombien and Hoole (2013), and corroborated by studies examining the pre-speech babble of infants in a number of languages (Eilers, Oller, & Benito-Garcia, 1984; Enstrom, 1982). Essentially, the longer, slower constriction for velar consonants, as compared to labials, translates to a higher Bernoulli force that prolongs the constriction and leads to a higher build-up of air pressure behind the tongue back, which when released translates to a larger and more intense burst, and longer VOT. For Spanish stops, Williams (1977), Rosner et al. (2000) and Gibson, Fernández Planas, Gafos, and Ramirez (2015) found significantly longer lag times for velars than for labials (velar > labial). Importantly, all of these VOT patterns provide robust audibly recoverable cues from which to extract phonological information over the course of phonological development.

The type of transition between the articulatory gestures in clusters may also provide crucial perceptual cues that learners may use to form early syllabic templates. For example, a recent kinematic study of Spanish clusters (Gibson et al., 2017) using electromagnetic articulography addressing, among other things, the temporal arrangement of onset clusters, reveals the existence of a lag time (i.e. interplateau interval, IPI) between the two gestures in complex onsets (see Fig. 1). This transition serves as the articulatory base of what others have described as an intrusive vocoid-like segment between the consonants, which is audibly perceptible to some degree (Ramírez, 2006). Fig. 1 provides kinematic support for this particular timing relation between the gestures in Spanish clusters, where shaded rectangles represent the articulatory constriction of the different articulators in a word initial /kl/ cluster. The temporal distance between these rectangles (represented as empty space between (b) and (c)) is the IPI:

In addition to the global timing pattern (open transition) between gestures in C1C2 Spanish clusters, a segmental effect was also found based on the C2 (Gibson et al., 2017). In Spanish C1C2 clusters, the C2 can be either a lateral or tap. Taps are typified by a retraction of the tongue back that facilitates an increase in velocity needed for the tongue tip to reach its alveolar target in a run-and-hit (in Catford's words) manner (Catford, 1988). This extra movement of the tongue back preceding the apical constriction translates to longer IPIs in stop-tap clusters as compared to stop-lateral clusters.