Journal of Memory and Language 96 (2017) 138–154

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Novel phonotactic learning: Tracking syllable-position and co-occurrence constraints

Amélie Bernard *

McGill University, Centre for Research on Brain, Language & Music, Canada

ARTICLE INFO

Article history:
Received 16 October 2015
Revision received 27 May 2017
Available online 21 June 2017

Keywords:
Adult language acquisition
Phonotactic learning
Statistical learning
Syllable position
Local co-occurrences
False recognition

ABSTRACT

Are syllable-level and co-occurrence representations simultaneously available when one learns novel phonotactics? After training on word-medial consonant restrictions (e.g., word-medial onsets P/Z, codas D/F, and cross-syllable consonant clusters FP/DZ in items like baF.Pev, tiD.Zek), adults falsely recognized novel items containing restricted consonants with the same co-occurrences (e.g., FP) more often than those with different co-occurrences (e.g., FZ) when syllable-position information was kept constant (e.g., vuF.Pet vs. vuF.Zet; Experiment 1). Thus, adults tracked co-occurrence information. Additionally, even when co-occurrence information was different from training, participants more often falsely recognized novel items that contained restricted consonants in the same (e.g., onset-Z) rather than different syllable positions (e.g., coda-Z), whether the restricted consonants were in the same (word-medial, e.g., vuF.Zet vs. vuZ.Fet, Experiment 2) or different word positions (word-edge, e.g., Zut.veF vs. Fut.veZ, Experiment 3). Thus, adults also tracked syllable-level information. These findings show that adults spontaneously represent sound sequences at multiple levels.

Introduction

For an English speaker, *dap* seems like a better potential word than *dah* (with a fully pronounced 'h' at the end); yet, *dah* is a possible word in Persian (meaning 'ten'). What makes English speakers, but not Persian speakers, reject *dah* as a potential word is the fact that it violates how sounds can be arranged (phonotactics) in English, but not in Persian. Knowledge of phonotactic constraints affects speech processing: facilitating speech segmentation (e.g., Mattys & Jusczyk, 2001; Mattys, Jusczyk, Luce, & Morgan, 1999), speech repetition (e.g., Munson, 2001; Vitevitch & Luce, 2005), and novel word learning (e.g., Graf Estes, Edwards, & Saffran, 2011; MacKenzie, Curtin, & Graham, 2012; Storkel, 2001; Vitevitch, Armbrüster, & Chu, 2004). Sensitivity to phonotactic constraints is exhibited early (e.g., Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Luce, & Charles-Luce, 1994; Zamuner & Kharlamov, 2016 for recent review) and remains flexible, allowing, in some cases, adult second-language learners to be nearly as sensitive to the phonotactics of their second language as to those of their first language, and nearly as sensitive as native speakers of that language (Weber & Cutler, 2006). Moreover, adults quickly learn novel phonotactic constraints such as P-starts and F-ends words from brief auditory (e.g., Bernard, 2015; Chambers, Onishi, & Fisher, 2010; Onishi, Chambers, & Fisher, 2002) or production experience (e.g., Dell, Reed, Adams, & Meyer, 2000; Goldrick, 2004; Goldrick & Larson, 2008; Kittredge & Dell, 2011; Warker & Dell, 2006, 2015; Warker, Dell, Whalen, & Gereg, 2008; Warker, Xu, Dell, & Fisher, 2009), enabling them to respond differentially to novel words like *paf* and *fap* that either follow or violate these constraints, as demonstrated by their rates of false recognition, repetition latencies, or production accuracies.

The fact that phonotactic knowledge leads to enhanced speech processing enables us to use phonotactic learning and generalization as a means to better understand how humans represent speech sounds and sound sequences. Evidence about how phonotactics are represented is mixed and somewhat ambiguous, in part because phonotactic learning has been studied from within different subdomains of speech processing, and thus has been described with variable terminology (for a similar discussion see Rapp, Buchwald, & Goldrick, 2014), and in part because relevant factors have often not been controlled. For example, some studies have shown a processing advantage for items with high, rather than low, phonotactic probability (e.g., Jusczyk et al., 1994; Vitevitch & Luce, 2005; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997), with phonotactic probability commonly operationalized as a combination of syllable position and a type of local co-occurrence

* Address: Department of Psychology, 1205 Dr Penfield, Montreal, QC H3A 1B1, Canada.

E-mail address: amelie.bernard2@mail.mcgill.ca

http://dx.doi.org/10.1016/j.jml.2017.05.006

information, making it impossible to know which information source participants used. In one study, English-learning infants preferred items like *riss* to items like *yowdge*, where *riss* contains both (1) sounds that occur frequently in their respective positions in English (i.e., in consonant-vowel-consonant words, R is frequently the onset or syllable start, I is frequently the vowel, and S is frequently the coda or syllable end) and (2) biphones that are frequent in English words (i.e., R followed by I, and I followed by S are combinations of sounds that often occur within English words), and *yowdge* contains both (1) sounds that occur rarely in their respective positions and (2) biphones that are rare (Jusczyk et al., 1994). However, since syllable-position and co-occurrence information were conflated in these studies, it remains unclear whether participants represented phonotactics at the level of the syllable, at the level of co-occurrences, or both. These studies are thus uninformative with regard to whether English speakers prefer *dap* over *dah*, because they have represented the fact that, in English, P is a better coda than H, or the fact that, in English, A and P form a better combination of sounds than A and H.

Since syllable-position and co-occurrences have often been conflated, one can ask whether people are sensitive to ***syllable-position constraints*** alone (when co-occurrence information is held constant; e.g., H can be an onset, but cannot be a coda in English). This idea is similar to what has been called *positional segment frequency* (e.g., Jusczyk et al., 1994; Vitevitch et al., 1997), *language-wide constraints* (for restrictions present in the native language), *experiment-wide constraints* (for restrictions present in an artificial language; e.g., Dell et al., 2000; Warker & Dell, 2006, 2015), or *consonant-position constraints* (e.g., Chambers, Onishi, & Fisher, 2003, 2011; Chambers et al., 2010; Onishi et al., 2002) in that representing syllable-position constraints involves tracking the position in which segments occur relative to a syllable-sized unit.

Similarly, one can ask whether people are sensitive to ***local co-occurrence constraints*** alone (when syllable-position information is held constant; e.g., A and P co-occur more often than A and H). The concept of local co-occurrence is akin, but not identical, to the concept of *biphones/diphones* (e.g., Gathercole, Frankish, Pickering, & Peaker, 1999; Majerus, Poncelet, Van der Linden, & Weekes, 2008; Munson, 2001; Richtsmeier, Gerken, Goffman, & Hogan, 2009; Richtsmeier, Gerken, & Ohala, 2009, 2011; Vitevitch et al., 1997; Weber & Cutler, 2006) in that it involves tracking sounds that are adjacent to one another. Unlike biphones/diphones, however, local co-occurrences could, in theory, be represented without linear information (i.e., AP and PA are, at some level, equivalent in that both are combinations of adjacent A and P) while biphone/diphone frequency calculations have tended to encode linear order (e.g., AP would have a different biphone frequency than PA). Thus, in some sense, biphones/diphones could be considered types of co-occurrences, that is, co-occurrences with order information. The concept of local co-occurrence is also related to the concept of *transitional probability* (e.g., Aslin, Saffran, & Newport, 1998; Pelucchi, Hay, & Saffran, 2009a, 2009b; Saffran, Aslin, & Newport, 1996) as it involves tracking the occurrence of adjacent units. However, in the case of local co-occurrences (as investigated here), the unit of interest is the phoneme, while in work on transitional probabilities, the unit of interest tends to be larger (i.e., the syllable).

Linguistic theories of phonological encoding are split with regard to the role of the syllable (see also Goldrick, 2004 for similar ideas); while some linguistic theories argue that phonotactic constraints are represented with reference to the syllable, which is decomposable into onset, nucleus, and coda constituents (e.g., Blevins, 1995), others argue that syllables are actually not sufficient and/or necessary for phonotactic encoding (e.g., Hirsch, 2014; Steriade, 1999). Nevertheless, previous psycholinguistic data hint at the ability to represent phonotactics at the level of the syllable. Adults appear to be sensitive to sound constraints at the level of the syllable in artificial-language learning experiments in which particular consonants were restricted to the beginning or the end of one-syllable consonant-vowel-consonant words (e.g., Chambers et al., 2003; Goldrick, 2004; Goldrick & Larson, 2008; Kittredge & Dell, 2011; Onishi et al., 2002; Warker & Dell, 2006, 2015; Warker et al., 2009). English speakers, for example, were more likely to accidentally produce *meNG* (where the critical sound, NG, is capitalized and represents a single consonant, /ŋ/) than *meH* (where the critical sound, H, is capitalized and represents a fully produced /h/) when attempting to say *mek*, showing sensitivity to the syllable-position constraints of English (e.g., in English, the sound NG (as in *siNG*) but not H (as in *Horse*) occurs at the end of syllables; called 'language-wide constraints' in Dell et al., 2000). Furthermore, after only brief training on one-syllable words that started with S and ended with F, participants were more likely to accidentally produce *keF* than *keS*, thus showing a sensitivity to experimentally established syllable-position constraints (e.g., the sound F, but not S, occurs at the end of syllables; called 'experiment-wide constraints' in Dell et al., 2000). Moreover, participants' sensitivity to experimentally established constraints was comparable to their sensitivity to the native-language constraints (Dell et al., 2000). Thus, adults were able to learn novel phonotactics that could have been represented at the syllable level (e.g., F is coda). Results were similar even when local co-occurrence information differed between training and test. Specifically, after training on words like *saF* and *paF*, adults were faster to repeat (and infants preferred to listen to) words like *peF* than *Fep* even though local co-occurrence information was novel (F co-occurring with E rather than with A) in all test items (Chambers et al., 2010, 2011), suggesting that phonotactics may be represented at the syllable level and abstracted away from the co-occurrences in which they were learned.

However, because determining the role of the syllable in phonotactic learning was not the goal of their research, neither Dell et al. (2000) nor Chambers et al. (2010, 2011), manipulated syllable-position directly. Because they trained and tested participants on a single 1-syllable word structure (CVC), it was not necessary for participants to represent the constraints at the level of the syllable to successfully learn the regularities: representational levels other than position within the syllable (e.g., relative to linear ordering or relative to word edges/silence) could have accounted for adults' success in learning the regularities. In CVCs, all codas are also in the third linear position, in the final slot of the word, followed by silence, etc. Thus, the constraint on F from *saF* and *paF* in training could have been represented at the syllable-position level ('F is a syllable coda'), enhancing processing for *peF* but not *Fep*. But the same regularity could also, however, have been represented (and extended to *peF* but not *Fep*) relative to the linear ordering ('F comes third'), relative to silence ('F is before silence'), or relative to the word edge ('F ends words'). Thus, while the above-mentioned results suggest that phonotactic constraints can be represented at the level of the syllable, they are not definitive. Work using different syllable or different word structures is needed to directly investigate whether syllable-level representations are spontaneously accessed during learning of novel phonotactic patterns.

Although there are fewer studies of artificial-language learning using stimuli other than CVCs, the ones that are available tend to be consistent with the idea of syllable-level representations. For instance, Bernard (2015) found support for syllable-level learning while directly manipulating syllable position by varying word structure and word position between training and test words. For example, participants trained on items like *buF.Pak* (where the period indicates the syllable boundary and capital letters indicate restricted consonants) distinguished between test items such as