Perception, production, and individual differences

doi:10.1017/S0142716418000176

Werker provides a compelling story for how infants' speech perceptual abilities develop and change over the first years of life. Less well understood is the relationship between perception and production abilities. Given pervasive evidence of individual differences in production, this raises many questions about possible individual differences in perception as well, and how this might influence the course of language acquisition more generally. This Commentary reviews some of the types of individual differences found in production, some of what we know about how these are linked to perception, and areas for further research.

It has long been known that children follow certain general patterns when acquiring their first words. However, there are also individual differences, in the timing of *when* children begin to produce their first words as well as in the actual phonological *shape* that these early words take. Thus, many children use their first words by 18 months, with a word spurt around 50 words, reaching 300 words by 24 months (cf., Fenson, Dale, Reznick, Bates, & Thal, 1994). However, others start to use their first words as early as 10 or 11 months, and are using short sentences by 18 months (e.g., two of the children in the Providence Corpus; Demuth, Culbertson, & Alter, 2006). Others ("late talkers") may only begin to use their first words after the age of 2, catching up around the age of 3 (cf., Rescorla, Roberts, & Dahlsgaard, 1997). All the above is considered to be within the normal range for children with no obvious evidence of language disorders. Yet little is known about the possible differences in the timing or characteristics of children's perceptual abilities, and if or how this might forecast these differences in production. Thus, although there are reports of perceptual sensitive at 6 months correlating with global measures of vocabulary size 1 or 2 years later (e.g., Tsao, Liu, & Kuhl, 2004), very little is known about the relationship between the perception and production of the same contrasts. For example, despite the fact that many children's first words take the form of simple, "unmarked" segmental, syllable, and prosodic word structures (e.g., Demuth, 1995; Gnanadesikan, 2004), we know little about why this evolves differently depending on the child, even for a given language.

Thus, the field of child phonology, including children's early word productions, appears to differ from that of the field of infant speech perception. However, this may be an artifact of the methods used. Studies of children's phonological development have often taken the form of case studies, examining the speech productions of a few children, often longitudinally, providing an in-depth look at individual developmental patterns over time (e.g., Brown, 1973). In contrast, infant speech perception studies have tended to involve groups of infants, reporting the development of language-specific perception abilities at a certain age. Although some theories of perception explicitly appeal to articulatory

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press

gestures in making claims about perception (e.g., Best's, 1995, perceptual assimilation model), there has been little explicit research exploring the relationship between perception and production abilities in the same children. This is partly due to the different ages typically investigated, with much of the field of infant speech perception focusing on infants below the age of 2, and much of the child phonology literature focusing on children from the age of 2 onward. This, plus the different nature of the online versus speech production methods typically used, which necessitate the study of groups versus individuals, respectively, has led to little cross-talk between the two fields. Fortunately, this situation is slowly beginning to change, with concomitant advances in our theoretical understanding of the important connections between early perception and production.

In the following, I review some of what is known about individual differences in production, at the segmental level, interactions between segments and syllables, and at the level of prosodic words and prosodic morphology. I then briefly discuss recent research from our lab that begins to explore individual differences at the perception/production interface, discuss the possible mechanisms behind these individual differences, and identify areas for further research.

INDIVIDUAL DIFFERENCES IN THE PRODUCTION OF SEGMENTS, SYLLABLES, AND PROSODIC WORDS

Jakobson (1941) was one of the first to suggest that the segments children first use would be those that maximized contrasts. Thus, he proposed that the low vowel /a/ might be the first vowel to appear, along with the obstruent sonorant contrast /p/ versus /m/. Although the validity of Jakobson's proposal has never been confirmed, many of these or similar segments do seem to be some of the first to appear in children's early babbling and words, with consonants that are harder to produce, such as fricatives, affricates, and liquids later acquired (cf. Smit, 1993). Thus, children's early productions tend to include the more "unmarked" segments that occur widely across many languages. Of course, different languages have different phonological systems, and this also influences when and how phonemic contrasts are learned. For example, recent research on the early acquisition of Fulfulde (a West-Atlantic Niger–Congo language spoken in Mali and widely throughout West Africa) shows that, before 11 months, at least some children are using complex implosives and prenasal stops, and that implosives (e.g., /ɓ/) account for more than 37% of the obstruents one child produced (cf. Cissé, Demolin, & Vallée, 2011). It is not yet known if this was a particularly precocious child, or if this is a general pattern found for many Fulfulde-speaking children. It would then be interesting to know how and when perceptual sensitivity to, for example, implosive–nonimplosive contrasts emerges for such children.

Other more "marked" segments can also appear early in children's speech. Dental clicks are produced early in languages like Xhosa (Gxilishe, 2004), whereas palatal alveolar clicks only appear later, and in Sesotho-speaking children's words around the age of 3 (Demuth, 2003). One might expect perceptual

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press

sensitivity to all such consonants to appear long before the articulatory demands of production are met. The same might be thought to be true for segments like fricatives and liquids, where the articulatory challenges of producing such forms delays their appearance in children's speech (cf. Lin & Demuth, 2015, for ultrasound study of English "l" development from 3–7 years).

Individual differences have long been documented at the segmental level in production, with some children preferring certain segments over others (Schwartz & Leaonard, 1982). For example, some have been shown to produce (many) words with labial–coronal sequences, or labial first words (cf. Yavaş, 1995), depending on the language. Word truncation processes can also vary from child to child, where different segments may be preserved. When truncating trisyllabic words like *banana* to a disyllable, many children omit the first syllable altogether, producing a form like *nana*. Others, however, use the form *bana*, preserving the less sonorant (less marked) consonant in the onset position (cf. Demuth, 1996a; Pater, 1997). Markedness in the selection of coda consonants also shows individual differences, with some English-speaking children's first coda consonants being composed of less marked (more sonorant) codas (such as fricatives), whereas many others first use the most frequent (but more marked) coda consonant /t/ (Stites, Demuth, & Kirk, 2004).

There are also many individual differences in the use of syllable structures and how these are produced. Some children begin using coda consonants with their first words, but many do not, producing predominantly reduplicated or consonant–vowel–consonant–vowel (CVCV) word forms for their first words (Demuth, 1995). A few, apparently precocious, children then go through a brief early stage of development (around 16 months) where they produce heavily released final (coda) consonants that can result in an "epenthetic" vowel, especially following voiced target codas (cf. Demuth et al., 2006; Vihman & Velleman 2000). With respect to complex onsets, some children truncate word initial *s*-clusters by consistently dropping the /s/ (e.g., *stop ~ top, slip ~ lip*), whereas others preserve the more sonorant consonant, resulting in *sop* for *stop*, but *lip* for *slip* (cf. papers in Goad & Rose, 2003). Individual differences in the use of consonant clusters are also found in Dutch-speaking children's early words, with some showing earlier use of clusters in the onset (CCVC), but others first showing complexity in the coda (CVCC; Levelt, Schiller, & Levelt, 2000). Thus, there are many well-attested individual differences in both the timing and the realization of segments between children learning the same language. It is not clear if or how this might relate to perceptual differences as well. Little is known about individual differences in the perceptual sensitivity to such segmental/syllabic contrasts, and how this might relate to these individual differences in production.

Individual differences are also found in the form of children's early prosodic words. For example, some Dutch-speaking children shift the placement of stress on words with a weak–strong (unstressed–stressed, iambic) stress pattern to a strong–weak (stressed–unstressed, trochaic) stress pattern, in keeping with the dominant stress pattern of the language (cf. Demuth, 1996b; Fikkert, 1994). Although there has been much research on infant sensitivity to language-specific stress patterns (e.g., Jusczyk, Cutler, & Redanz, 1993; Yeung & Nazzi, 2014),

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press

little is known about individual differences in perception, and how this may relate to these differences in production.

Some children also tend to use "proto function words" often known as "fillers" (e.g., *a wan a more* "I want some more"), whereas others do not (cf. Peters, 1983). It is unclear why this is the case, and if children's perceptual sensitivities to function words (e.g., Shi, Werker, & Cutler, 2006) show concomitant individual differences. Some Spanish-speaking children actively truncate lexical items even while producing early "proto-articles" (*amueka* "la muñeca" (the doll)), while others never truncate lexical forms, only producing articles with fully formed words (cf. Demuth, Patrolia, Song, & Masapollo, 2012). Individual differences are also found in English-speaking children's use of articles, with most children prosodifying articles with the proceeding monosyllabic word (e.g., *He [likes the] dolly*; Gerken, 1996), but some actually stressing the article, initially producing it as its own prosodic word around the age of 20 months (*He [likes] [the] dolly*; cf. Demuth & McCullough, 2009). Again, it is likely that perception of such forms is robust (cf. Gerken & McIntosh, 1993), but some of our resent research below suggests a close connection between perception and production.

INDIVIDUAL DIFFERENCES AT THE PERCEPTION/PRODUCTION INTERFACE

It is unclear if any of the individual differences in production discussed above correlate with differences in perception. However, there are indications that some of them might. For example, we have found that, in spontaneous speech, 2-year-olds tend to produce more inflectional morphemes, such as 3rd person singular *–s*, when these occur utterance finally compared to utterance medially (Song, Sundara, & Demuth, 2009). The final syllables of a phrase undergo phrase-final lengthening (Klatt, 1976; Lehiste, 1973); we thus thought that phrase-final contexts might provided a longer durational context for all segments (and morphemes) to be produced. However, this phrase-final lengthening might also enhance the perceptual salience of these morphemes, making them easier to perceive as well. This was tested and found to be the case: both 20- and 27-month-olds showed sensitivity to 3rd person singular inflectional morphemes utterance finally, but not utterance medially, suggesting a close connection between perception and production (Sundara, Demuth, & Kuhl, 2011). This finding also suggests that higher perceptual salience may facilitate "learning," and this will then also be reflected earlier in production.

However, most verbs in English occur utterance medially, raising questions about when and how children begin to perceive and produce inflected verbs in this position. Further investigation revealed that those infants who showed higher 3rd person singular production abilities in an elicited imitation task also showed greater sensitivity to the presence versus (ungrammatical) absence of the morpheme utterance medially (e.g., *She cries/cry now*) during a listening task (Sundara et al., 2011). Such findings suggest a tight correlation between perception and production at the level of the individual.

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press

What could be the mechanism underlying these individual differences? Werker suggests that perceptual sensitivity interacts with meaning and referential context. We know that individual children's receptive and expressive vocabulary sizes show large individual differences from the time these can be measured. Perhaps this is a result of more efficient perceptual development, leading to earlier production as well? Recent findings suggest that this may be the case. In some of our recent research exploring sensitivity to plural morphology during a listening-while-looking task, we found that 24-month-olds with higher vocabularies were sensitive to plural /s/ (as in *cats*; though not /z/, as in *dogs*; Davies, Xu Rattanasone, & Demuth, 2017). Further investigation confirmed that /s/ was longer in duration than /z/, and that this probably leads to its greater perceptual salience (despite the more frequent occurrence of /z/ as a plural morpheme). Thus, children who produce more words seem to be able to tune in to the greater perceptual salience of this /s/ allophone, attending to its meaning as a morphological unit and encoding it into their emerging grammar prior to doing the same with the voiced plural allomorph /z/.

Such findings suggest that there is a good correlation between the development of perceptual abilities and production, at least in global terms. However, identifying such correlations requires collecting multiple pieces of information from the same participants. For young infants, information on receptive vocabulary and/or the use of communicative gestures may provide a means for exploring such individual differences. Werker suggests that perceptual sensitivity develops as part of a multisensory system, where attention to multimodal (facial) cues may also play a role in learning how to articulate certain sounds. Perhaps children who attend to multimodal cues exhibit earlier onset of speech, and/or more targetlike use of early segments and prosodic word structures.

Recent research does suggest that motor and language development are closely linked (e.g., Iverson, 2010). This is further supported by research on infants fitted with cochlear implants, which finds that receptive communication abilities and fine motor skills before implantation, along with early gesture use, are significant predictors of vocabulary size at 12–15 months postimplantation (cf. Bavin, Sarant, Prendergast, & Leigh, 2017). This suggests that early motor abilities may correlate well with individual abilities in production, and perhaps perception as well.

SUMMARY

In sum, recent findings suggest that there may be a closer connection between perception and production abilities than often thought, with individual differences emerging in both. This can be illuminated by collecting multiple language/gesture/cognitive measures on the same children, shedding light on the nature of individual differences (cf. Cristia, Seidl, Junge, Soderstrom, & Hagoort, 2014). Such approaches are becoming more common in many other fields of psycholinguistic research (e.g., Montrul & Tanner, 2017), and will hopefully become more common in infant/toddler research in years to come.

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press

ACKNOWLEDGMENTS

This research was supported by Macquarie University, the Australian Research Council (ARC) Centre of Excellence in Cognition and Its Disorders Grant CE110001021 (awarded to Crain et al.), and ARC Grant FL130100014 (to K.D.).

REFERENCES

Bavin, E. L., Sarant, J., Prendergast, L., & Leigh, G. (2017). *Predictors of early vocabulary development for children with CIs*. Paper presented at the 14 Triennial meeting of the International Association for the Study of Child Language, Lyon, France, July 17–21, 2017.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange & J. J. Jenkins (Eds.), *Cross-language speech perception* (pp. 171–204). Timonium, MD: York Press.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Cissé, I., Demolin, D., & Vallée, N. (2011). *The acquisition of plosives and implosives by a Fulfulde-speaking child aged from 5 to 10;29 months*. Paper presented at the 17th International Congress of Phonetic Sciences, Hong Kong, August 17–21, 2011.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Development*, *85*, 1330–1345.

Davies, B., Xu Rattanasone, N., & Demuth, K. (2017). Two-year-olds' sensitivity to inflectional plural morphology: Allomorphic effects. *Language, Learning and Development*, *13*, 38–53.

Demuth, K. (1995). Markedness and the development of prosodic structure. In J. Beckman (Ed.), *Proceedings of the North East Linguistic Society* (Vol. 25, pp. 13–25). Amherst, MA: University of Massachusetts, GLSA.

Demuth, K. (1996a). The prosodic structure of early words. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 171–184). Mahwah, NJ: Erlbaum.

Demuth, K. (1996b). Alignment, stress and parsing in early phonological words. In B. Bernhardt, J. Gilbert, & D. Ingram (Eds.), *Proceedings of the International Conference on Phonological Acquisition* (pp. 113–124). Somerville, MA: Cascadilla Press.

Demuth, K. (2003). The acquisition of Bantu languages. In D. Nurse & G. Phillipson (Eds.), *The Bantu languages* (pp. 209–222). London: Routledge.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech*, *49*, 137–174.

Demuth, K., & McCullough, E. (2009). The prosodic (re)organization of children's early English articles. *Journal of Child Language*, *36*, 173–200.

Demuth, K., Patrolia, M., Song, J. Y., & Masapollo, M. (2012). The development of articles in children's early Spanish: Prosodic interactions between lexical and grammatical form. *First Language*, *32*, 17–37.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., & Thal, D. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*, (Serial No. 242).

Fikkert, P. (1994). *On the acquisition of prosodic structure* (PhD dissertation. University of Leiden).

Gerken, L. A. (1996). Prosodic structure in young children's language production. *Language*, *72*, 683–712.

Gerken, L. A., & McIntosh, B. (1993). The interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*, 448–457.

Gnanadesikan, A. (2004). Markedness and faithfulness constraints in child phonology. In R. Kager, J. Pater & W. Zonneveld (Eds.), *Constraints in phonological acquisition* (pp. 73–109). Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press

Goad, H., & Rose, Y. (Eds.). (2003). Segmental-prosodic interaction in phonological development: A comparative investigation. *Special issue of Canadian Journal of Linguistics*, 48, 139–452.
Gxilishe, S. (2004). The acquisition of clicks by Xhosa-speaking children. *Per Linguam*, *20*, 1–12.
Iverson, J. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language*, *37*, 229–261.
Jakobson, R. (1941). *Kindersprache, aphasie und allgemeine lautgesetze*. Uppsala: Almqvist & Wiksells boktryckeri a.-b.
Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675–687.
Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.
Lehiste, I. (1973). Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America*, *54*, 1228–1234.
Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of syllable types. *Language Acquisition*, *8*, 237–264.
Lin, S., & Demuth, K. (2015). Children's acquisition of English onset and coda /l/: Articulatory evidence. *Journal of Speech, Language, and Hearing Research*, *58*, 13–27.
Montrul, S., & Tanner, D. (2017). Individual differences and retrieval interference in L2 processing. *Bilingualism: Language and Cognition.* Advance online publication. doi: 10.1017/S136672891600095X
Pater, J. (1997). Minimal violation and phonological development. *Language Acquisition*, *6*, 201–253.
Peters, A. (1983). *The units of language acquisition: Monographs in applied psycholinguistics*. Cambridge: Cambridge University Press.
Rescorla, J., Roberts, J., & Dahlsgaard, K. (1997). Late talkers at 2: Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, *40*, 556–566.
Schwartz, R., & Leonard, L. (1982). Do children pick and choose? An examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language*, *9*, 319–336.
Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*, 187–198.
Smit, A. B. (1993). Phonologic error distributions in the Iowa-Nebraska articulation norms project: Consonant singletons. *Journal of Speech and Hearing Research*, *36*, 533–547.
Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular *-s*. *Journal of Speech, Language, and Hearing Research*, *52*, 623–642.
Stites, J., Demuth, K., & Kirk, C. (2004). Markedness versus frequency effects in coda acquisition. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (pp. 565–576). Somerville, MA: Cascadilla Press.
Sundara, M., Demuth, K., & Kuhl, P. (2011). Sentence-position effects on children's perception and production of English 3rd person singular *–s*. *Journal of Speech, Language, and Hearing Research*, *54*, 55–71.
Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*, 1067–1084.
Vihman, M. M., & Velleman, S. L. (2000). The construction of a first phonology. *Phonetica*, *57*, 255–266.
Yavaş, M. (1995). Phonological selectivity in the first fifty words of a bilingual child. *Language and Speech*, *38*, 189–202.
Yeung, H. H., & Nazzi, T. (2014). Object labeling influences infant phonetic learning and generalization. *Cognition*, *132*, 151–163.

Katherine Demuth
Macquarie University

https://doi.org/10.1017/S0142716418000176 Published online by Cambridge University Press