

The reality of phonological forms: a reply to Port

Carol A. Fowler

Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA

Abstract

I suggest four grounds on which an argument can be made that phonological language forms are not merely emergent properties of the public language use of members of a language community. They are: (1) the existence of spontaneous errors of speech production in which whole consonants or vowels misorder or are replaced; (2) the necessary existence of language “particles” used by individual language users in order for words to be able to be coined; (3) the remarkable effectiveness of alphabetic writing systems and the tight coupling among skilled readers of orthographic and phonological language forms; (4) the finding that, by late infancy, children have discovered phonological constancies despite phonetic variation.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

I agree with many of the observations that Bob Port makes in his paper (Port, 2010). We disagree about what some of them mean. First, I think that we agree about what a linguistic theory reflects; it reflects what works for a language community. It is not what sits inside a language module in someone’s mind or brain. It is, on average, what language users know, but is not what any of them knows in particular. It is what emerges from public language use by members of a language community. We also agree that phonological components of the language are not abstract symbols in the mind. Finally, we agree that literacy affects how some of us think about the spoken language. Literacy may even affect how we know our spoken language.

Here is where I think we disagree. My belief is that consonants and vowels are real properties of the living language – the language used by members of a language community. Moreover, I see them necessarily as something that every language user has knowledge of and uses. However, they are not something in memory; they are linguistic actions of the vocal tract. Of course, to use language forms in talking requires having knowledge that supports their use. If episodic memory systems of the sort that Port describes can do that, then fine. If they cannot, then some other kind of cognitive system must do so.

There are four kinds of evidence that I will review that make the case for me. I will discuss each in turn; finally, I will discuss specificity effects.

First are spontaneous errors of speech production. Talkers make errors showing that something like phonetic segments are in their repertoire (e.g., the Reverend William Spooner’s alleged “mardon me padam” for

E-mail address: carol.fowler@haskins.yale.edu

“pardon me madam”). Second is what Abler (1989) and Studdert-Kennedy (2003) refer to as the “particulate principle of self-diversifying systems”. To have lexicons of indefinitely large size, language communities need a small set of transmutable segments, consonants and vowels. I believe that, because individuals, not communities coin words, individuals must have knowledge of and use these particles. Third is the rather astonishing success of alphabetic writing systems. Fourth is recent evidence by Best and colleagues that toddlers at 19 months of age, but not at 15 months of age show sensitivity to phonological constancies among words produced in different dialects. Those constancies must be present if they are possible to find.

2. Speech errors

Regarding “but in the spoken language, there is no evidence of any alphabet at all”. (p. 10): That is not so. Speakers make errors. Spontaneous errors of speech production are errors that speakers do not ordinarily make, but do on occasion. Shattuck-Hufnagel (1983) reported that errors involving single segments predominated in her collection of sublexical errors. Among exchange errors (e.g., *sprit blain* for *split brain*) in which the error unit could be unambiguously identified, whole vowel or consonant errors constituted 66% of the total. This estimate is challenged somewhat by findings on errors collected in the laboratory (e.g., Goldstein et al., 2007) where the speakers’ articulatory gestures can be observed. Many errors transcribed as whole-segment errors may be more complex, involving blends of intended and intruding speech gestures. However, in my view it is very likely that pure instances of whole-segment errors (such as *beef needle soup* for intended *beef noodle soup*; Dell, 1986) do occur. They are reported to pattern very much like whole word errors where the units participating in errors are indisputable (*We have a laboratory in our own computer* for *We have a computer in our own laboratory*; Fromkin, 1973) in consisting of movement errors (perseverations as in *beef needle soup*, anticipations as in *take a nook at the news*, and exchanges: *mardon me padam*); and noncontextual substitutions (e.g., *pick someone out* for intended *kick someone out*). These errors strongly suggest that speakers have as units for speech production something very much like the phonetic or phonemic segments proposed to be speech units by linguists. Further, because these errors seem to occur in *planning* for speech production, they imply to me that, in the cognitive system that engages in speech planning, there are discrete transmutable representations of consonants and vowels.

Two studies of preliterate children (Jaeger, 2005; Stemberger, 1989) suggest that they produce whole segment errors (although, in Stemberger’s data, children made a higher proportion of feature errors than did adults). If that is right, learning to read does not create particulation into consonant and vowel units.

3. The particulate principle of self-diversifying systems

Abler (1989) proposed that there is a small number of systems, including genetic recombination, chemical compounding and spoken languages that are “self-diversifying”. That is, they are generative systems that create variation without bound. An important feature of all of the systems is that they are “particulate”. That is, their components are discrete particles that combine without blending.

Contrast this with baking. When we make a cake, we combine ingredients such as flour, sugar, eggs, and butter. When the ingredients are combined, they lose their individual identities. No identifiable flour, no identifiable eggs, etc. Rather, there is cake batter. But now imagine combining the phones, /ae/, /k/ and /t/. We can have *act*, *tack*, and *cat*, three different words composed of the same phones. Blending of components does not occur. It is this particulate property that allows indefinitely large lexicons of distinct word forms in spoken language.

If I understand what Port wrote in his paper, he may agree that the emergent language of a language community has particles, but he does not agree that individual members of language communities do. But it is individual members of language communities that coin words, composing them of particles.

4. The success of alphabetic writing systems

Port writes that our impression that words are composed of discrete ordered consonants and vowels is likely a consequence of our being literate in an alphabetic writing system. However several considerations suggest to me that our impression must be based on more than that.

First, there are alphabetic writing systems. Their inventors must have had the impression that the spoken language had units to which the letters would correspond. Yet they had no alphabetic writing system to give them that impression.

Second, there is the remarkable success of alphabetic writing systems. That is, most people taught to read become very skilled. Reading becomes sufficiently effortless that they read for pleasure. Is it likely that these writing systems would be so effective if they mapped onto language forms that individual members of the language community did not command?

Finally, there is considerable evidence that skilled readers of a great variety of writing systems access phonological information within a very short time after seeing print (see, e.g., [Frost, 1998](#), for a review). What, in memory, are they accessing?

5. Sensitivity to phonological constancies

[Millikan \(2003\)](#) notes that words have “lineages”; that is, histories of usages of which individuals intercept a subset. The lineages that we intercept let us know how a word is used; what it means. “Cat”, for example, is used to refer to a particular type of animal. But, of course, utterances of /kaet/ all are different. They differ in dialect, in the particular idiosyncracies of the speaker, in speaking rate and more. How do we intercept a lineage; that is, how do we know that token instances of a word are instances of the same word? Port’s answer seems to be that we do not. But we do.

[Best and colleagues \(2009\)](#) have begun to explore when young children notice phonological constancy despite phonetic variability. She finds that, by 19 months of age, but not by 15 months, children recognize the sameness of words when they are spoken in very different dialects of English (in the study, Connecticut vs Jamaican English) despite the phonetic variation that distinguishes the dialects. The (American) children showed a familiarity preference, that is a willingness to listen longer to familiar than to unfamiliar words. For the 19 month olds, familiar words spoken in an unfamiliar dialect counted as familiar words.

So there is some sameness that listeners detect. What might it be? Port rejects the assumption of linguists and psychologists that their phonetic alphabet will eventually be found to have physical reality. He writes: “But physical definitions for segments and features have never been found. To succeed would require that some acoustic property be found everywhere that a [d] or [i] or a [-voice] feature occurs”. (p. 6) But no it wouldn’t require that. The acoustic signal is not where phones and features live. They are linguistic actions of a vocal tract (e.g., [Browman and Goldstein, 1992](#)), and, there, gestural properties can be found that are always there for [d] (a tongue tip constriction gesture at the alveolar ridge of the palate), [i] (a palatal constriction of the tongue body), and [-voice] (vocal fold abduction).

6. Specificity effects

Port cites evidence that, when language users hear words, they retain not just information that identifies the words, but also information that identifies the speakers of those words. Moreover, linguistic and nonlinguistic information remain bound together in a common episodic memory trace. This has been interpreted (e.g., by [Goldinger, 1998](#)) as evidence that the mental lexicon is not a memory of abstract word types, but a memory of episodic encounters with word tokens. I do not understand why, for Port, such specificity effects provide information that memory for words is absent memory for familiar language units such as phones. In the particular theory of lexical memory that he cites ([Goldinger, 1998](#)) listeners retain information about spoken words in vectors of features. However, in the vector, phonetic features are represented, and voice features are represented separately from the phonetic features. So why cannot such findings preserve the notion of phonetic particles in speech separate from speaker-specific information?

Acknowledgement

Preparation of the manuscript was supported by grants HD-01994 and DC-02717 to Haskins Laboratories.

References

- Abler, W., 1989. On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures* 12, 1–13.
- Best, C.T., Tyler, M.D., Gooding, T.N., Orlando, C.B., Quann, C.A., 2009. Development of phonological constancy. *Psychological Science* 20, 539–542.
- Browman, C., Goldstein, L., 1992. Articulatory phonology: an overview. *Phonetica* 49, 155–180.
- Dell, G., 1986. A spreading activation theory of retrieval in sentence production. *Psychological Review* 93, 283–321.
- Fromkin, V., 1973. *Speech Errors as Linguistic Evidence*. Mouton, The Hague.
- Frost, R., 1998. Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin* 123, 71–99.
- Goldinger, S., 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251–279.
- Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., Byrd, D., 2007. Dynamic action units slip in speech production errors. *Cognition: International Journal of Cognitive Science* 103, 386–412.
- Jaeger, J., 2005. Kids' slips. Lawrence Erlbaum Associates, Inc, Mahwah, NJ.
- Millikan, R., 2003. Defense of public language. In: Antony, L., Hornstein, N. (Eds.), *Chomsky and His Critics*. Blackwell, Malden, MA, pp. 215–237.
- Port, R.F., 2010. Rich memory and distributed phonology. *Language Sciences* 32, 43–55.
- Shattuck-Hufnagel, S., 1983. Sublexical units and suprasegmental structure in speech production planning. In: MacNeilage, P. (Ed.), *The Production of Speech*. Springer-Verlag, New York, pp. 109–136.
- Stemberger, J., 1989. Speech errors in early child language production. *Journal of Memory and Language* 28, 164–188.
- Studdert-Kennedy, M., 2003. Evolutionary implications of the particulate principle: imitation and the dissociation of phonetic form from semantic function. In: Knight, C., Studdert-Kennedy, M., Hurford, J. (Eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press, Cambridge, pp. 161–176.