Journal of Phonetics 53 (2015) 125–140

Contents lists available at ScienceDirect

Journal of Phonetics

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

Creating the cognitive form of phonological units: The speech sound correspondence problem in infancy could be solved by mirrored vocal interactions rather than by imitation

Piers Messum[a,*], Ian S. Howard[b]

[a] Pronunciation Science Ltd, 112 Warner Road, London SE5 9HQ, UK

[b] Centre for Robotics and Neural Systems, School of Computing, Electronics and Mathematics, Plymouth University, Plymouth PL4 8AA, UK

ARTICLE INFO

Article history:
Received 20 December 2014
Received in revised form
28 June 2015
Accepted 27 August 2015
Available online 2 October 2015

Keywords:
Phonological units
Underlying representation of speech
Speech acquisition
Correspondence problem
Development of pronunciation
Imitation
Mirroring

ABSTRACT

Theories about the cognitive nature of phonological units have been constrained by the assumption that young children solve the correspondence problem for speech sounds by imitation, whether by an auditory- or gesture-based matching to target process. Imitation on the part of the child implies that he makes a comparison within one of these domains, which is presumed to be the modality of the underlying representation of speech sounds. However, there is no evidence that the correspondence problem is solved in this way. Instead we argue that the child can solve it through the mirroring behaviour of his caregivers within imitative interactions and that this mechanism is more consistent with the developmental data. The underlying representation formed by mirroring is intrinsically perceptuo-motor. It is created by the association of a vocal action performed by the child and the reformulation of this into an L1 speech token that he hears in return. Our account of how production and perception develop incorporating this mechanism explains some longstanding problems in speech and reconciles data from psychology and neuroscience.

1. Introduction

This special issue of the Journal of Phonetics focuses on how phonology – the science of speech sound systems – can contribute to the longstanding debate about the nature of speech units in the human brain. Phonology, though, is the outcome of practical processes: a child learning to comprehend and pronounce a particular language. Basic assumptions about how this takes place set the terms of the debate. In particular, in the learning of pronunciation it has always been assumed that the child solves the correspondence problem for speech sounds by imitation. That is, the child uses his[1] own judgement of similarity between what he recovers from the speech input (an acoustic pattern in some theories or a pattern of gestures in others) and what he produces in return to match this. This judgement informs and improves his subsequent production in a 'matching to target' process. A judgement of similarity is only possible between images that are comparable, so the child is assumed to be operating either with auditory or motor primitives, and the favoured one of these is then considered to be the form for the underlying representation of speech in the human brain.

In our work, we have been investigating how the pronunciation of L1[2] is learned. For a number of reasons we argue that it is unlikely that children solve the correspondence problem by acoustic imitation of caregiver speech. Instead we suggest it is plausible that they find a solution in the dynamics of caregiver–infant interaction. Here, imitation takes place, but it is usually the caregiver imitating the child, rather than vice versa. In the gestation period of speech, the form of the imitation is rarely simple mimicry; instead a caregiver reformulates her child's output into L1, giving him evidence of the correspondences between what he does and what she considers its linguistic significance to be.

* Corresponding author. Tel: +44 20 7274 6306.
E-mail addresses: p.messum@pronsci.com (P. Messum), ian.howard@plymouth.ac.uk (I.S. Howard).

[1] We avoid continual use of "he or she", "his or her", etc., by using pronouns which describe interactions between a female caregiver and a male child.

[2] In this article we use the following abbreviations: L1 for 'first language', ME for'mirrored equivalence', SBE for 'similarity based equivalence', AS for 'awareness of sensation', MP for 'meaningful perception', IM for 'inverse model', PM for 'perceptuo-motor' and VMS for 'vocal motor scheme'.

http://dx.doi.org/10.1016/j.wocn.2015.08.005

Like Moulin-Frier, Diard, Schwartz, and Bessière (2015), we have used a computational model to investigate some aspects of this matter. Our computational agent, Elija (Howard & Messum, 2007, 2011, 2014), models speech acquisition by children. To examine the development of pronunciation, we aimed to endow him with capacities for production, perception and cognitive activity that are no greater than those of a human infant. With these, he was able to develop a repertoire of potential speech sounds, to interact with human subjects taking the role of his caregiver, to learn speech sound correspondences from these interactions, and finally to use these correspondences to learn the pronunciation of simple words in the language of each caregiver.

In this article, we describe the theoretical background to Elija, starting with the issue of how pronunciation is learnt. We discuss previous proposals for how children might solve the correspondence problem, which are principally acoustic matching theories. We describe these as examples of a 'Similarity Based Equivalence' (SBE) mechanism. We then explain how well-attested mirroring behaviour seen on the part of caregivers supports an alternative proposal which we describe as a 'Mirrored Equivalence' (ME) mechanism. We argue that at critical stages of child speech development ME provides a better explanation of observed phenomena that any SBE account. ME would generate an intrinsically perceptuo-motor cognitive form for phonological units and would support a new account of how speech production and perception develop in a child, which we describe.

2. Learning to pronounce

2.1. How the pronunciation of words is learned

To explain our account of the development of pronunciation in a child, we need to distinguish the activities of learning how to pronounce particular words from learning how to pronounce speech sounds. The first of these, the mature skill of learning the pronunciation of a new L1 word, is readily accessible to introspection and is uncontroversial: the speaker parses the word he has just heard into a string of speech sounds and says, in his own voice and in the same order, a string of speech sounds that he knows his listeners will take to be equivalent to the ones he heard (see Fig. 1). A speech sound in this context is a syllable, or even a couple of syllables, formed of one or more phonemes (Guenther, Ghosh, & Tourville, 2006, p. 283) which occurs commonly enough to form part of Levelt's 'mental syllabary' (Levelt, Roelofs, & Meyer, 1999, p. 5). In order to achieve this ability to imitate, the speaker must first learn how to produce speech sounds that will be taken to be equivalent to the ones he hears. To characterise these two distinct learning activities, we use the terminology 'learning to pronounce a word' and 'learning to pronounce' respectively.

It takes several years for a child's pronunciation to approach an adult level of competence (Dodd et al., 2003). Thus learning to pronounce to the point of mastery must represent a significant practical challenge for a child, even though it has been taken to be conceptually straightforward. The general assumption has been that learning to pronounce is a self-supervised process of auditory 'matching to target': having identified a speech sound, the child tries to copy what he hears and then judges for himself the similarity of his output to the target. He uses this to improve his subsequent attempts.

Within this account, 'perception' has been taken as a precursor to the development of production. Therefore it has been assumed that a single phonological lexicon begins to develop as the child's mind grapples with the perceptual data presented to it, and that this lexicon goes on to inform production. As a consequence, scholarly interest has focussed on the question of how phonology is acquired during perception, rather than on the apparently secondary process of how pronunciation is actively learned. However, as we show below, if speech sounds are not learnt by imitation then it is possible for an output phonological lexicon both to develop and to be structured independently from a perceptual lexicon. Thus the learning of the pronunciation of speech sounds is an issue to be addressed in understanding the genesis of phonology.

2.2. The correspondence problem for speech sounds

For a mature speaker to learn to pronounce a word by parsing its component speech sounds and reproducing them using his own voice, he first needs to discover how to produce speech sounds that will be taken to be equivalent to the ones he hears. This requires

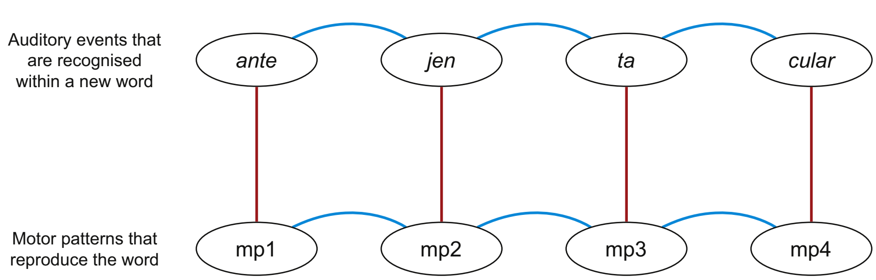

Fig. 1. The mature skill of learning the pronunciation of a new word requires (1) learning the identity and ordering of the speech sounds heard, and (2) prior to this, learning the 'vertical' links between speech sounds heard and the motor patterns that can be used to reproduce them. *source*: When a word is heard for the first time, the speaker parses it into speech sound elements. For example, he may decompose 'antejentacular' into 'ante – jen – ta – cular'. He can reproduce these four auditory events using four motor patterns, each of whose output he knows will be taken by his listeners to be equivalent to what he has heard. Thus he learns to pronounce the word by serial imitation. However, he must have previously learnt the 'vertical' links between speech sounds he hears and their corresponding motor patterns. The correspondence problem for speech sounds is the question of how he achieves this: either using some form of imitation or by some other mechanism. Fig. 1 and terminology adapted from Heyes (2001).

Forbidden

You don't have permission to access this resource.

Apache/2.4.25 (Debian) Server at daneshyari.com Port 443