Language Acquisition in Human Infants

Can Control Theory Provide an Integrative Account?

Emma-Jane Greig

Warren Mansell

University of Manchester, UK

Address for correspondence:

Dr Warren Mansell

Senior Lecturer in Psychology
School of Psychological Sciences
Coupland I
University of Manchester
Oxford Road
Manchester M13 9PL
Email: warren.mansell@manchester.ac.uk
Tel: +44 (0) 161 275 8589
Website: http://www.psych-sci.manchester.ac.uk/staff/131406

Abstract

We revisit the phenomena of language acquisition and propose that a control theory perspective has the capacity to integrate existing accounts. Learning language involves the use of sensory feedback by the infant to achieve its goals via its primary caregiver and to eventually become an active individual within society. The infant learns increasingly sophisticated methods of reproducing linguistic sounds, organised in a hierarchical structure. This learning process involves the random variation and selection of sounds (babbling) that are reproduced faithfully despite variations in musculature, auditory environment and the listener. This account shares features with Guenther's (1995) DIVA model, Lindblom's (1989) H & H theory and Moore's (2007) PRESENCE model of speech. We explain how a control theory model has the capacity to integrate these approaches, model them, and further incorporate features such as the use of metaphor, mental rehearsal, symbolism and the role of language in culture.

Language Acquisition in Human Infants

Can Control Theory Provide an Integrative Account?

An accurate understanding of how children develop language can pave the way to confronting larger issues concerning the function of language within the individual and throughout society. It is clear that understanding, learning and the use of appropriate language enables individuals to establish their own identity, relationships and a form of self expression. This is why effective communication development has become increasingly fundamental in achieving fluid interactions between society members and gaining their acceptance (Syder, 1992). Indeed, language has been considered an 'unwritten passport' to an individual's societal belonging (Johannson, 2005).

Our article provides an overview of the challenges in developing such a comprehensive account of language acquisition and assesses whether a perspective based on Perceptual Control Theory (PCT; Powers, Clark, & McFarland, 1960; Powers, 1973, 2005) can integrate several existing approaches (Guenther, 1995; Lindblom, 1989; Moore, 2007b).

Language Acquisition – The Core Phenomena to Explain

From birth, human infants are bombarded with communication, sound stimulation and sensations. Their world almost instantly becomes dominated by the desire to communicate and express themselves verbally (Barnes & Mercer, 2003; Syder, 1992; Van Riper & Emerick, 1990). Acquiring a language involves learning its sounds, sound patterns and the way in which words can be combined. Each of thousands of languages in the world is based on a complex system of rules for combining different elements. Typically these combinations are represented as a hierarchy (Siegler, Deloache, & Eisenberg, 2006). For example, complex syntactic structures of sentences such as "*when the driver saw the red light, he stopped the car*" can be broken down to reveal combinations of sound units called phonemes. At these lower levels, the phonemes approximating "c" and "ar", for example, would be combined to produce the word "*car*" and so forth. A range of evidence supports such a hierarchical account of language, including studies of computational modelling and neurobiology (Kiebel, Daunizeau, & Friston, 2009; Poeppel, Idsardi, & van Wassenhove, 2008).

The first step in children's language learning therefore involves grasping an understanding of these elementary units of sounds. The next progression then leads to combining individual phonemes, such as *ar* or *ee,* to make meaningful semantic units. For example, *ar* would then become c*ar*t and *ee* could be grouped to make sw*ee*t. These small units of meaning, familiar as words, are also called *morphemes* (Deutsch, 1981; Edwards & Shriberg, 1983; Siegler, et al, 2006).

From here, going upward in a hierarchical organisation, the child then needs to learn the rules of syntax. Essentially, these refer to the rules that dictate the order in which words can be arranged in a clause. Once syntax has been established, the fourth and final step is termed *pragmatic development* (Siegler et al., 2006). This refers to a complex understanding of how language should be used i.e. in which context it might be appropriate to use the past tense, or on a more specific level, that the word "please" often precedes a request or demand. The pragmatic and contextual nature of language appears to be ubiquitous within adult speech (Leudar, Sharrock, Hayes, & Truckle, 2008), and therefore requires an explanation within an account of language acquisition

There are a number of observations from the first few months of life that provide clues about the process of language acquisition. Early on, babies appear to be preparing for speech production by making a variety of sounds. At around seven months of age infants begin to *babble*. Essentially, this is a random coupling of vowels and consonants. Infants whose hearing is intact will produce sounds which follow a distinctive pattern i.e. a consonant followed by a vowel *"baba"*. As the infants' babbling becomes more varied, it begins to mimic rhythmic and intonational patterns of everyday language (Siegler et al., 2006). The importance of this process is clear. For example, studies of deaf children indicate that auditory feedback from the infant's own speech appears to be necessary for babbling to commence in the first year of life (Koopmans-van Beinum, Clement, & van den Dikkenberg-Pot, 2001). Thus, not only the generation of babbling, but the child's perceptual feedback of her own babbling appear to be relevant to language acquisition.

There is also evidence that the caregiver's communication style towards the child is important within the first year. Mothers of infants tend to speak in 'motherese', a style of intonation that emphasises certain fundamental verbal sounds over others. There is evidence that infants have a preference for this style of speech (Cooper, Abraham, Berman, & Staska, 1997), and that it allows infants to orient for longer to important grammatical units of speech that they need to master (Nelson, Hirsch-Pasek, Juscyk, & Cassidy, 1989). Thus, it appears that certain properties of a caregiver's speech improve language acquisition by targeting the developmental needs of the infant.

The production of recognizable words begins at around one year of age. By the end of the infant's second year, most infants are quite capable of producing short sentences. Fully formed pragmatic understanding is seen to become active at around five years of age when a child has acquired a vast amount of oral language including verbal labels for concepts of size and colour, question forms and tense markers (Syder, 1992).

In summary, language is thought to be built up in a serious of stages i.e. phonemes > morphemes > syntax > pragmatics, represented in a hierarchy. Importantly, the majority of children move through the same stages of language learning and communication development and at roughly the same age (Gerber & Kraat, 1992). It is this logical progression of language acquisition that has become the basis for language production and acquisition research. Research has branched away from simplistic observational coding experiments to specifically begin to track stages of speech development, explore the importance of speech feedback, social interaction and child-directed speech, and most recently develop the use of computational speech modelling technology (Guenther, 1995; Hofe & Moore, 2008; Henke, 1966; Lindblom, 1989; Perkell, 1980). In the next section we introduce the framework of Perceptual Control Theory in order to illustrate how its core principles of negative feedback, hierarchical organisation and learning through 'reorganisation' provide a firm framework on which to understand and model language acquisition (PCT; Powers, Clark, & McFarland, 1960; Powers, 1973, 2005).

Perceptual Control Theory (PCT)

PCT was originally formulated by William T. Powers and colleagues in 1960, and has since developed into an integrative theoretical account of human behaviour (Powers, Clark, & MacFarland, 1960; Powers, 1973, 2005). It has led to the development of self-regulation theory within psychology (Carver & Scheier, 1998), Affect Control Theory within sociology (Heise, 1977), and a range of applications including artificial intelligence and robotics (Powers, 1979), education (Cziko, 1992) and political psychology (d'Agostino, 1995). PCT is related to an earlier approach to understanding human behaviour called *cybernetics* (Wiener, 1948). The foundation behind this suggests that people are essentially intricate control mechanisms, whose control involves the maintenance of numerous intrinsic variables at internally selected values (e.g. Ashby, 1952). Within PCT, there are four key principles of human functioning and behaviour; *control, hierarchical organisation, conflict* and *reorganisation,* each of which will be outlined below.

Control

The basic idea in PCT is that all intentional actions are performed so as to bring the values of perceptual signals closer to their reference values. In other words, 'behaviour is the control of perception' (Powers, 1973); a wide variation in observable behaviour – in this case the complex form of spoken language – can achieve a single purpose – communication between individuals. We have different perceptions of how we feel our experience of something should be. For example, we have a standard for how warm we want to be and for how much food we can eat until we feel satisfied. This standard is also known as a *reference value*; life is regarded as a constant process of comparing how things are with how we want things to be. If the two fail to match the system acts to reduce this discrepancy. Thus, control is a process of reducing *error* and getting levels as close as possible to the desired reference point (Powers, 1973; 2005).

Fundamentally, PCT is based on a negative feedback loop framework (see Figure 1). There are four main functions within this framework: *reference value*, *input, comparator,* and an *output function*.

- *The input function* (represented as (i) in Figure 1) is responsible for sensing disturbances in the environment.
- *Reference value* (r) is the internal standard that the organism is striving to achieve.
- The *comparator function* (c) detects any discrepancy present that deviates from the internal reference value.
- Lastly, the *output function* (o) is essentially process generating the behaviour that is performed to reduce the discrepancy

----- INSERT FIGURE ONE AROUND HERE ----

Importantly, the negative feedback loop allows research to establish operational computerised and mathematical models of control by living systems (Marken & Powers,1989; McPhail, Powers & Tucker, 1992). This approach paves the way to explore new areas of fine motor control, such as language acquisition and its development.

Hierarchical Organisation

According to Powers (1973), control systems are organised in a hierarchy. Importantly within PCT, it is only the lowest level that uses behavioural output to control perception i.e. the appropriate muscle contractions in the tongue, mouth and vocal tract that produce the desired string of sounds. The lower-level systems, in other words, do not know anything about movements or relationships. They only sense and control position (Powers, 1988). The loops arranged above them, higher up in the hierarchy, set the reference values or the goals for the loops below. Thus, higher order systems develop for the purpose of regulating and balancing lower order systems in the living thing (Powers, 1990) i.e. they select the reference values for the lower levels.

Within a PCT hierarchy, the lowest level controls intensities, for example loudness of sound and muscle tension. The perceptions at this level are the signals generated by sensory nerve endings. As the intensity of stimulation increases, the frequency of firing of sensory nerves increases. At the next level up, intensities are organised into *sensations*. Sensations are linked directly to elementary experiences such as colour, taste and force. In terms of infant speech development, the change in intensity of a sound at level-one forms a recognisable sound quality. It is suggested that phonemes form organised patterns of sound sensations and are therefore perceived at a third level of *configurations*.

There are proposed to be a total of 11 levels in the hierarchy (Intensity < Sensation < Configuration < Transitions < Events < Relationships < Categories < Sequences < Programs < Principles < System Concepts). However, it is acknowledged that this is open to modification following empirical enquiry (Powers, 1988). Importantly, the highest levels allow more complex organisations of perception such that complete *programs* of action can be controlled. Programme actions have been compared to a store of possible contingencies. By this we mean having an established network of possibilities i.e. if condition A holds, take branch 1; otherwise take branch 2 (Powers, 1988). These individually stored programmes form the components of *principles,* which are higher level abstractions and have social and moral properties. In terms of language development, it is the principle level that holds information concerning the unwritten rules of language and social interaction i.e. turn-taking phenomenon and appropriate cultural etiquette. These *principles* are located at the highest level within a system (e.g. the self, the workplace) that the individual is striving to uphold.

It is proposed that the infant is born with volitional control of only the lowest hierarchical level – the intensities of her experience, and that the higher levels of perceptual control develop in stages over the first two years of development (Plooij & Rijt-Plooij, 1990). As a child grows and develops, neuron connections are established through contact with external sensations (Plooij & Rijt-Plooij, 1990). Alongside this, levels of perceptual control are adjusted and refined. Thus, the stage-like nature of child development observed by a wide range of researchers is regarded as the development of an increasingly sophisticated hierarchy of control (Plooij & Rijt-Plooij, 1990). Essentially, as a child develops a new level of control within the hierarchy, complexity and ability increase. This is mirrored in patterns of language development mentioned earlier (phonemes > morphemes > syntax > pragmatics). Table 1 (adapted from Powers, 1998) attempts to clarify the ages and behavioural outputs associated with the development of linguistic control systems. Behavioural shifts have been noted in children at 2 months, 7, 12 and 18-21 months (Plooij & Rijt-Plooij, 1990). These appear to coincide with significant milestones of control system levels as depicted below.

---- INSERT TABLE 1 AROUND HERE ----

Reorganisation

From a control systems view, the hierarchical structure is built up stepwise through major neural reorganisations in the nervous system at specific ages, which allows the human infant to perceive a new type of invariance (Plooij & Rijt-Plooij, 1990; Powers, 1973). *Reorganisation* is essentially a random-walk learning process (e.g. Marken & Powers, 1989), which is where parallels have been drawn to neuron connections as seen in initial brain development. Simplistically, if we consider a newborn infant's desire to vocalise and attract their caregiver's attention, there is no previously set reference value that explicitly tells the infant how much to contract his facial muscles or which direction to move his tongue in order to produce a sufficient sound. This needs to occur as a consequence of a system sending random signals for tensing muscles arbitrarily (Syder, 1992). Within PCT, it is the properties of the control systems that change by a random amount which in turn has consequences on the neural signals. These properties include the *weightings* at which signals are sent to, and received from, the adjacent lower level. At the lowest level in the control hierarchy, nerve signals do not dictate reference values, but the level of contraction of muscles. Thus, in the case of speech, the random variation in signal would produce random muscle contractions of the intercostal and facial muscles. These signals only stop when the infant experiences the desired perception (e.g. the infant produces an appropriate mapping of a sound that attracts its caregiver's attention) (Powers, 1990). Two aspects of language acquisition described earlier are clear from this account: that it is important for the infant to receive ongoing perceptual feedback from her own utterances, and that the presence of a caregiver is critical.

Within a PCT model of speech development, babbling would form the majority of random signals being sent out via nerve impulses, which in turn contribute to sound perceptions and articulatory goals. Therefore, in terms of 'learning' when an infant encounters a new signal pattern (e.g. the appropriate tongue muscle contraction in order to produce the letter 'd') it is then stored in memory within the control unit. When this is the case, then it has become a controllable 'perception' (Powers, 1973).

If, in this example, the infants fails to produce the 'd' sound due to inappropriate tensing of the tongue muscle, but she wants to restore the experienced sound perception, the higher system must signal the unit in the control hierarchy where that perception is stored, in order to reset it as a reference signal for the systems moving the tongue and associated vocal apparatus. Hence, the role of reorganisation is to create the learned control hierarchy. Problems occur when the infant continually encounters problems producing a specific sound. This may be caused by *conflict.*

Conflict

According to PCT, conflict occurs when two control systems attempt to control the same quantity, each of which have different reference values. On a more serious level, and perhaps beyond the scope of the present paper, long-term conflict has been linked to the development of psychopathology including anxiety, depression and psychosis (Mansell, 2005; Powers, 1973). In the majority of cases, acquiring language is based on short-term conflict. A key cause of both long-term and short-term conflicts is *arbitrary control* (Mansell, 2005; Powers, 1973). Arbitrary control exists whereby executing one goal will conflict with a second goal the individual also wishes to achieve (Higginson, Mansell, & Wood, in press; Powers, 1973). For example, in the earlier example of attracting the attention of a caregiver, one goal might be that the infant wants a soft toy that is out of reach. The problem is that the infant lacks a sophisticated vocabulary to communicate this desire. Instead, the infant must choose between getting immediate attention through crying, and the alternative goal that demands time spent reorganising reference frames for known letter sounds to produce the word 'toy', or words similar to that affect. Progress is then made when the infant resolves the conflict. Essentially, this resolution comes from reorganisation of the system, as described earlier.

Reorganisation therefore has two functions within the control process. Firstly, it forms the basic reference values for syllables and sounds in the initial stages of babbling. Secondly, it also operates via trial-and-error processes and randomly alters the way we manage our goals until the conflict is reduced. For example, the infant may find an alternative way of attracting its caregiver's attention in order to receive the toy, or she may just change his priorities so that vocalising intelligible speech is not as important to her as getting his caregiver's attention through crying or random babbling.

Fundamental to the resolution of conflict is awareness. Within PCT, for reorganisation to be effective in the long term, awareness must be directed to the higher- level goals. For example, an infant may learn that generating morphemes through combinations of phonemes is more likely to attract positive attention from the caregiver than generating phonemes that do not form recognisable words.

The crucial thing to note is that experiencing and resolving conflict is integral to an infants' development. An infant who does not resolve conflict ultimately does not progress to the next stage of development. The levels of control within the hierarchy are prohibited from continuing to the next level upwards. It is proposed to be an underlying factor in language delay or other known speech sound disorders.

The Complete System

It is now possible to summarise the complete system in operation. In the initial stages of speech development (approximately three months old), an infant may have an articulatory goal of producing the speech sound *da*. However, before this combined reference value is established, random signals displayed through babbling will have taken place. The outcome of this will have created vague but distinguishable reference values for the sounds 'd' and 'a'.

Pronouncing the 'd' sound requires the operation of internally set reference values that function to achieve the intended sound. An example of this might include the appropriate muscle contraction of the tongue to move it into contact with the soft palate of the mouth, or the level of control exerted to the inter-costal muscles and diaphragm for speech to occur (Syder, 1992). If the desired sound is produced, this is set as a perception within the infant's control hierarchy. Producing the *da* sound may elicit a desired reaction from the parent, e.g. she receives a tickle. If the infant wishes to recreate the desired sound because of the newly formed positive association, a higher order system must signal to the place in the control hierarchy where the perception is stored.

Once activated, this perception serves as a reference signal for the lower systems in charge of movement of facial muscles. Once a child has learnt the intended reference value for the *da* sound, negative feedback enables the child to monitor internal control (Robertson & Powers, 1990). This is achieved through auditory feedback and internal proprioception i.e. sensed tongue position.

Focusing on the dashed line shown in Figure 2, if a child fails to produce the 'd' sound, he will experience increased error. This is represented as (1) on Figure 2. The resulting error then drives the system in a direction that minimizes the difference between the desired perceptual signal (reference) and the input (Powers, 1973) i.e. altering tongue contraction rate. If the error signal is prolonged, meaning that the child constantly mispronounces the'd' sound, reorganisation will result (see Figure 3, (2)). The consequence of this results in the child changing the parameters in order to reach its desired articulatory goal of *'da'*.

---- INSERT FIGURE 2 AROUND HERE ----

This reorganisation process will cease only when it results in restoration of the intrinsic error signal to zero, meaning the discrepancy between values (reference value and input) reduces and a new reference value for "d" is set (Runkel, 2003). The amount of produced error is then reduced, as shown in Figure 2, Section (3), until the infant encounters the next challenge in language development. In other words, no further reorganisation is needed because the amount of spoken error has been reduced and the infant has resolved the encountered difficulty. Importantly, Figure 2 shows that language learning is a continual stepwise progression. Infants must experience error in order to progress to the next stage or achieve a new level of control. Consequentially, language learning would not occur unless there is error (Runkel, 2003).

The filled-in line (shown in Figure 2), represents the behaviour of the negative feedback loop *without* reorganisation. One should note that although there is some reduction in the amount of error produced due to negative feedback received from outputs, the overall amount of error left unresolved remains much higher than when reorganisation occurs. Moreover, the amount of produced error appears to oscillate over time. If left unaddressed, this continued error would represent impaired language acquisition.

The complexity of this process may help to illustrate why the initial stages of language acquisition can prove difficult for young infants. Learning weightings for individual letter sounds and then letter blends is a complex and exhausting task (Van Riper & Emerick, 1990). However, once an infant has grasped the basics and a typical reference value has been set, development to word blends and finally whole word constructions can ensue.

More recently, speech modelling systems have dissected observable speech development processes and integrated them into testable and precise models of fine motor control (Moore, 2007). The outcome of this has enabled a greater depth of understanding of the stages involved in language acquisition and speech production (Moore, 2007c).

Speech Modelling

Early models of speech production proposed spatial targets for articulators (Henke, 1966) and muscle length targets (Cohen, Crossberg & Stark, 1988). Unfortunately, these models failed to include explanations of compensatory movements, e.g. if a conversation ensued whereby the speaker was talking to another individual whilst chewing some food. In the majority of cases a listener would be able to distinguish a large proportion of target words produced by the listener, despite the sounds not being an identical match to the original reference frame. This is a direct example of how a system can operate across a variety of different situations by using a "best-match" mechanism, but it was not incorporated into earlier models of speech production. Later models hypothesized abstract functions of vocal tract shape and speech signal output (Guenther, 1995; Lindblom, Lubker, & Gay, 1979; Perkell, 1980). However, again these theories failed to provide a fully computational account of how speech parameters are formed. In particular, many theories have overlooked how speech modelling and production directly link to neurosystem development.

Largely, speech modelling has addressed numerous key aspects described within PCT but has not integrated them into one comprehensible and testable theory. Three of these approaches are discussed below. These theories are Guenther's (1995) DIVA model, Lindblom's (1989) H & H theory and Moore's (2007b) Predictive Sensorimotor Control and Emulation (PRESENCE) model of speech. They are each largely based on target *ranges* rather than explicit target *positions* as suggested by earlier approaches. This allows for variation in speech production i.e. 'da' could be similar to 'ta', which is more fitting to everyday speech.

DIVA Model

Guenther's (1995) DIVA (Directions (in orsosensory space) Into Velocities of Articulators) recognises the four distinct reference frames that operate throughout speech production. These are: *acoustic, phonetic, orsosensory* (*somatosensory*) and *articulatory* (*motor) reference frames*. Babbling in the DIVA model is produced by inducing random movements of the speech articulators (Guenther, 1995). These produced movements are largely constrained by "training sequences" that appear to relate to the specific infant to current neuromotor development that is specific to the infant. If the model recognises a configuration corresponding to a known speech sound, it activates a corresponding cell responsible for that sound. This subsequently drives learning in the phonetic-to-orsosensory mapping. The 'model' then learns the appropriate association and learning results from its own production and listening to that of 'others' (Guenther, 1995). Essentially then, the model learns by establishing reference frames through random production of sounds, similar to that of PCT.

The DIVA model (Guenther, 1995) is a very detailed computational model that has shown promising steps to explaining single motor behaviour in speech production. However, the model does not address anticipatory mechanisms, in terms of how the listener might receive or interpret a spoken message. Ultimately, there are two reference frames that coincide with one another: one for one's own speech production, and another reference frame that imitates or reflects how the listener might receive that message (Moore, 2007). A model that addresses just this was proposed by Lindblom (1989), named the H & H theory.

H & H Theory

The H & H (hyper- and hypo- speech) Theory suggests that individuals reorganise articulatory gestures and acoustic patterns in order to convey the clearest message the first time they speak, consequentially minimising energetic cost (Hofe & Moore, 2008; Lindblom, 1989).

Lindblom (1989) argues that it is more advantageous to the individual if the relayed spoken message is produced only once. For example, it would not be effective to whisper "how are you?" to the listener if one were at a music concert and the background noise was extremely loud. As a result, Lindblom (1989) proposes that the speaker adapts their spoken message in accordance to how they predict the listener will receive the message, taking into account environmental disturbances. Corrections are only made when it is judged that the perceived clarity does not match that which was intended (Hofe & Moore, 2008).

Similarities between Lindblom's (1989) theory and PCT are present in two areas. Firstly, feedback from auditory channels and proprioception play a major role in governing speaker behaviour (Hofe & Moore, 2008). Secondly, with respect to corrections being made, using PCT terminology the system is reorganised, thereby allowing the speaker to reproduce speech that is more fitting to the environmental conditions present and the perceived state of the listener (Moore, 2007c).

Lindblom's (1989) H & H theory has provided speech modelling research with an influential general framework that others have built upon and developed. Unfortunately, the H &H model does not explain a specific reaction to particular influences (Hofe & Moore, 2008). Specifically, Lindblom (1989) does not explain how these initial references are formed and in what time-frame. Moreover, the theory does not incorporate the hierarchical nature of language, nor how control over different aspects of language is maintained (Siegler et al., 2006)

PRESENCE model

An extension of Lindblom's (1989) H & H theory, is Moore's (2007b) Predictive Sensorimotor Control and Emulation (PRESENCE) model of speech. It includes a hierarchical negative feedback system of the self, in addition to a model of the 'predicted' listener that is explicitly based on Perceptual Control Theory. Moore (2007b) states that speakers tune their performance according to communicative and situational demands surrounding them and that of the listener. In other words, the speaker also has a concept or reference value for their listener, predicting how they will receive the message produced first time. This is known as the speaker's anticipatory mechanism.

Importantly, Moore's (2007b) model of spoken language processing contains four levels of hierarchy. These levels begin with a basic primary route for motor behaviour, which drives motor action and emulation of possible actions to follow. A second layer deals with sensory input and feedback, which allows the system to recognise if the desired intention has been met (Moore, 2007b). The third level represents a feedback path dealing specifically with behaviour dealing with the 'self' and the perceived 'other'. Finally, the fourth layer deals with the complex task of interpreting the needs and intentions of others.

In this respect, PRESENCE is built on the basis of control loops and feedback. Importantly, Moore (2007a) recognises that there is a hierarchy within speech production, in that higher levels control not just a concept of one's own language use, but that of the listener.

The potential of PCT to go beyond existing accounts

All three of the above models have touched upon important aspects of PCT, but none have integrated the individual components such as feedback loops and anticipatory mechanisms into a testable model. Arguably, therefore, PCT can provide an integrative and testable account of language processing that brings together work from Guenther's (1995), Lindblom's (1989) and Moore's (2007) models. However, we propose that PCT exceeds the three previous theories in its ability to explain beyond the initial stages of speech production. While a complete working model based on PCT has not been developed, the following sections illustrate some of the promises of a PCT-based account.

Symbolism and Context

PCT may provide an account of how speech sounds and words are translated into symbolism and used within a pragmatic sense, as highlighted earlier. For example, when confronted with the word "orange" we know that the word has two meanings; it is both a type of fruit and a colour. For many, in the appropriate context, the image is brought to the speaker's mind (Powers, 1990). Hence, words can change their meanings as the context changes (Runkel, 2003). This phenomenon is well recognised in the field of linguistics, and often termed Gricean meaning. Ultimately, it expresses that what 'the words say' (i.e. the semantics of that word) and what the speaker means is brought together to form a meaningful context (Carston, 2002).

Applying this to our PCT hierarchical framework, the use and understanding of pragmatics emerges late in development. It can therefore be assumed that as language 'users' we not only develop appropriate flexible reference frames for identifying words, but also reference frames for contextual understanding. If for example you hear "you have a red light", it could mean that you are holding a red light bulb or even that you are facing a red light at a traffic light in you car. The child constantly has to reorganise their internal system, not only to produce appropriate speech but also to understand other speakers around her. An understanding of the rules of pragmatics does not usually appear until the child reaches eighteen months, which once again conveys the hierarchical nature of language.

Another example by McPhail and Tucker (1990) illustrates the importance of representing language at higher levels of abstraction. They provide an example from reading - the words you are reading now consist of, on average, three to seven letters each. You read the words on this page at approximately at 100 words per minute. If asked to recite 300-700 *letters* per minute you might struggle with the given task, but you can think, speak or read 100 *words* per minute. In other words, the letters are transferred from individual sensations into configurations, categories, and relationships that form the words and phrases that appear on the page in front of you (McPhail & Tucker, 1990). This is essentially how an infant builds their understanding of the human language. Moreover, this stepwise progression contributes to the internal hierarchical system that processes, manipulates and organises our understanding and deliverance of language.

In many ways, our lives appear to be run by the consequences of performing programmed manipulations of words (Powers, 1990). Powers (1990) uses the example of calling a policeman a "pig". The attached stigmatism is not a reflection of the appearance of the policeman (we hope!), but of their own distaste towards his profession. Again, words seem to differ significantly depending on the context they are used in (Runkel, 2003). As depicted in Table 1, children start to grasp an understanding of pragmatics usage at around 18 months. However it is not until the age of six that children are seen to actively use symbolic reference frames in everyday speech (Siegler, et al, 2006).

Fitting this into the framework of PCT, it states that the symbolic meanings of words presented at the *program* level can be translated and broken down into reference perceptions and then activated at lower levels in the hierarchy (Powers, 1990). Taking the earlier example, although the motor coordination for pronouncing the word "pig" uses a programmed set of reference frames independent of context, higher levels in the control hierarchy are then responsible for applying the correct context to the word or statement - the consequence of which means that the statement makes sense within the given context and is coherently understood by the listener (Runkel, 2003).

Imagination, Planning and Metaphor

An extension to the use of language within symbolism is the use of mental rehearsal, which combines imagination, symbolism and 'thinking'. One example might be a situation in which one needs to collect grocery shopping after work: one may need to mentally rehearse what fruit or vegetables to purchase. The images associated with specific words are conjured up in awareness despite a lack of vocalisation. One *imagines* the relationships between words but refrains from actually making the memory derived images into active reference signals for speech (Powers, 1990; Runkel, 2003). Therefore, the rehearsal process operates internally, repeating itself as much as the individual wishes and, importantly, without any verbal output (Lakoff & Johnson, 1980).

This is an area in which PCT could be developed further. Powers (1973) describes the workings of an *imagination mode*, through which higher levels of a control hierarchy can act 'as if' a perception is experienced without activating the lower order levels of the hierarchy. Powers (1973) explains that during the imagination mode perceptual signals would bypass the lower order systems within the hierarchy, ultimately looping continuously between the higher-order systems. In this sense, the imagination mode is an isolated control system that is temporarily disconnected to lower-order systems. Only when a programme is found that satisfies the individuals' principles do the lower-order systems get "switched" on and the first move is executed (Powers, 1990).

In terms of language, imagination forms an integral part to childhood and adulthood development. Further research into the positive effects of early use of imagination has on language acquisition could be explored in more depth. In particular, PCT would predict that the development of higher levels within the control hierarchy (i.e. of principles and system concepts) would be associated with increased use of imagination. Additionally, do children with more active imaginations go on to have better abstract or concrete thinking skills? Is this a consequence of operating higher order levels in the hierarchy from a young age?

A further important field of interest that utilises language could be important here - the use of metaphor. Metaphors involve a rich use of internally rehearsed mental imagery that bridge two different conceptual domains (Lakoff & Johnson, 1980). In terms of PCT, the use of metaphor relies heavily on an individuals' flexibility to control multiple meanings, or in this sense, reference frames. In terms of the rules of traditional language use i.e. appropriate context and semantics, metaphors appear to defy and challenge our instinctive understanding. The phrase "the world is our oyster" challenges what we know about the two concepts. For instance, we know that the world does not look like an oyster through geographic maps and satellites, and we also know that oysters often contain precious pearls. Broken down these two concepts do not make sense. It is only when we start to elaborate and make the links between the two that something concrete is established. For example, the Oyster may contain a pearl→

the pearl has great value → I will seek my fortune in the world→ I consider the world to be my Oyster → I work on the principle that the world contains a 'virtual' pearl and I will seek fortune in the world. Importantly, this process appears to be rehearsed in our imagination until the desired perception is experienced and met with a sense of understanding (see also Stott, Mansell, Salkovskis, Lavender, & Cartwright-Hatton, 2010).

Language and Culture

Theories of language development have traditionally been separated from theories of language in its cultural context (McPhail, et al, 1992). However, PCT is a psychological theory that lends itself well to a sociological context. For example, McPhail et al. (1992) have utilised computer models based on PCT to model the behaviour of crowds in social gatherings.

McPhail et al (1992) proposed that two or more individuals may adopt the reference signal offered by the third party. An example would be two employees doing the task that their boss has told them to do. In terms of culture, individuals typically adopt the appropriate language practices of their culture. Using a PCT perspective, McClelland (1994) explains how language operates as a form of *collective control* whereby a group of individuals who share similar reference frames for specific words and linguistic structures can act together with great power than any one individual. Constructive examples would include groups of professions, such as lawyers, medics and engineers, who share the same terminology. A problematic example would include racism, with its shared vocabulary, through which individuals can try to control other people of a different race (e.g. using terms denoting inferiority).

Thus, a further missing link for a theory of language development would incorporate the linguistic differences and nuances of different social groups, and how these are passed across generations. As a result, the symbolism of language forms a 'shared meaning', whilst establishing identity and acceptance in a community (McPhail & Tucker, 1990).

Differences between cultures and social groups could be linked to the constraints of that community's language. Newcomers must acquire the language and its conventional uses if they are to survive within the community (McPhail & Tucker, 1990). Importantly, this emphasis on acceptance would be seen as a purposeful attempt by adults to pass on their culture and language to their offspring in a controlled manner for a variety of reasons (e.g. to maintain identity; to be seen as successful).

The developmental stage of the child would have a key influence on what is learned – i.e. at very young ages where the child is reorganising lower levels of perception; the influence would be felt in the way that phonemes are combined into morphemes. At later stages, the influence would affect naming and subsequently grammar. These proposals are clearly preliminary and yet testable.

Conclusion and Summary

Perceptual Control Theory has been introduced as an explanatory framework for language acquisition. The key principles of the theory - control, hierarchical organisation, conflict and reorganisation - were explained alongside several contemporary models of fine motor control and speech production. With origins based in control engineering, PCT aims to provide a truly mechanistic account of how change and development in communicative behaviour occurs and progresses. As a 'meta-theoretical' theory, it has the capacity to guide developmental psychologists and speech and language therapists in positive directions. It could guide early interventions concerning infants who express difficulties or potential conflicts in language production. There is an impetus for developing more comprehensive models of language acquisition based on PCT that not only incorporate reference frames for the speaker and listener, but also begin to model the hierarchical properties of language, its pragmatic use, the role of symbolism, mental rehearsal, metaphor and the importance of language within culture. This is clearly a vast project, but one that PCT, with its broad application across the social sciences, may have the capacity to deliver.

References

Ashby, W. R. (1952). *Design for a Brain.* Chapman & Hall, London: UK.

Barnes, M., & Mercer, G. (2003). *Disability.* Cambridge; Key Concepts, Polity Press.

Carston, R. (2002). Linguistic meaning, Communicated meaning and cognitive pragmatics. *Mind and Language, 17*, 127-148.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior.* New York: Cambridge press.

Cohen, M. A., Grossberg, S., & Stork, D. G. (1988). Speech perception and production by a self-organizing neural network. In Y. C. Lee (Ed.), *Evolution, learning, cognition, and advanced architectures.* Hong Kong: World Scientific Publishers.

Cooper, R., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behaviour and Development, 20,* 477-488.

Cziko, G. A. (1992). Purposeful behaviour as the control of perception: Implications for educational research. *Educational Researcher, 21,* 10-18.

D'Agostino, B. (1995). Self-images of hawks and doves: A control systems model of militarism. *Political Psychology, 16,* 259-268.

Deutsch, W. (1981). *The Child's Construction of Language.* London: Academic Press Inc.

Edwards, M. L., & Shriberg, L. D. (1983). *Phonology: Applications in Communicative Disorders.* San Diego: College-Hill Press.

Gerber, S., & Kraat, A. (1992). Use of a developmental model of language acquisition: Applications to children using AAC systems. *Augmentatic and Alternative Communication, 8,* 19-32.

Guenther, F, H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102,* 594-621.

Heise, D. R. (1977). Social action as the control of affect. *Behavioral Science, 22,* 163-177.

Henke, W. L. (1966). Dynamic articulatory model of speech production using computer simulation. *Unpublished doctoral dissertation,* Massachusetts Institute of Technology, Cambridge, MA.

Higginson, S., Mansell, W., & Wood, A. (in press). An Integrative Mechanistic Account of Psychological Distress and Therapeutic Change: The Perceptual Control Theory Approach. *Clinical Psychology Review.*

Hofe, R., & Moore, R. (2008). Towards and investigation of speech energetics using 'AnTon': an animatronic model of a human tongue and vocal tract. *Connection Science, 20,* 319-336.

Johansson, S. (2005). *Origins of Language: Converging evidence in language and communication research.* John Benjamins Publishing.

Kiebel, S., Daunizeau, J., & Friston, K. (2009). Perception and hierarchical dynamics. *Frontiers in Neuroinformatics, 20,* 1-9.

Koopmans-van Beinum, F.J., Clement, C.J. & Dikkenberg-Pot, I. van den (2001). Babbling and the lack of auditory speech perception: a matter of coordination. *Developmental Science, 4,* 61-70.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by, Part 11: Humans as symbol-using creatures. In *The production of reality: Essays and readings on social interaction* (Ed. O'Brien, J), pp103-118. Pine Forge Press.

Lindblom, B. (1989). Explaining phonetic variation: A sketch of the H&H theory. In *Speech Production and Speech Modelling* (Eds. Hardcastle, W & Marchal, A), pp403-433. Kluwer Academic Publishers.

Lindblom, B., Lubker, J., & Gay, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics*, *7,* 147-161.

Mansell, W. (2005). Control Theory and Psychopathology: An integrative approach. *Psychology and Psychotherapy: Theory, Research and Practice,* 78, 141-178.

Marken, R. S., & Powers, W. T. (1989). Random-walk chemotaxis: Trial-and-error as a control process. *Behavioural Neuroscience, 103,* 1348-1355.

McClelland, K. (1994). Perceptual control and social power. *Sociological Perspectives, 37,* 461-496.

McPhail, C. & Tucker, C. (1990) Purposive collective action. *American Behavioral Scientist, 34,* 81-94.

McPhail, C., Powers, W., & Tucker, C. (1992). *Simulating individual and collective action in temporary gatherings*. *Social Science Computer Review, 10,* 1-28.

Moore, R. (2007a). Spoken Language Processing: Piecing it all together. *Speech Communication, 49,* 418-435.

Moore, R. (2007b) PRESENCE: A Human-Inspired Architecture for Speech-Based Human Machine Interaction. *IEEE Transactions on Computers, 56,* 1176-1187.

Moore, R. (2007c). Spoken Language Processing by machine. In *The Oxford Handbook of Psycholinguistics* (Eds. Gaskell, G., Altman, G., Bloom, P., Caramazza, A., & Levelt, P), pp723-738. Oxford University Press.

Nelson, K., Hirsch-Pasek, K., Jusczyk, P.W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16,* 55-68.

Perkell, J. S. (1980). Phonetic features and the physiology of speech production. In B. Butterworth (Ed.), *Language production, Volume 1: Speech and talk* (pp. 337-372). New York: Academic Press.

Plooij, F., & Rijt-Plooij, H. H. C, van de. (1990). Developmental Trasnitions as Successive Reorganizations of a Control Hierarchy. *American Behavioural Scientist, 34,* 67-80.

Poeppel, D., Idsardi, W., & van Wassenhove, V. (2008). *Speech perception at the interface of neurobiology and linguistics.* Philosohpical transitions of the Royal Society, B, 363, 1071-1086.

Powers, W. T. (1973). *Behaviour: the control of perception.* Oxford, UK: Aldine.

Powers, W. T. (1979). The nature of robots. *BYTE, 4,* 132-141.

Powers, W. T. (1988). Making Sense of Behavior: The Meaning of Control. New Canan, CT: Benchmark Publications.

Powers, W. T. (1990). *Living Control Systems: Selected Papers of William T. Powers.* Gravel Switch, KY: Control Systems Group.

Powers, W, T. (2005). *Behaviour: The Control of Perception.* (2nd ed), Benchmark Publications; New Canaan.

Powers, W. T., Clark, R. K., & McFarland, R. L. (1960). A general feedback theory of human behaviour. Part II. *Perceptual and Motor Skills,11,* 309–323.

Robertson, R. J., & Powers, W. T. (1990). *Introduction to Modern Psychology: The Control-Theory View.* The Control Systems Group Inc: Kentucky.

Runkel, P. J., (2003). People As Living Things: the psychology of perceptual control. Hayward, CA: Living Control Systems Publishing.

Siegler, R., Deloache, J., & Eisenberg, N. (2006). *How Children Develop (2nd ed).* New York: Worth Publishers.

Stott, R., Mansell, W., Salkovskis, P. M., Lavender, A., & Cartwright-Hatton, S.

(2010). An Oxford Guide to Metaphors in CBT: Building Cognitive Bridges. Oxford, UK: OUP.

Syder, D. (1992). *An Introduction to Communication Disorders.* London: Chapman & Hall.

Van Riper, C., & Emerick, L. (1990). *Speech Correction: An Introduction to Speech Pathology and Audiology.* Englewood Cliffs, NJ: Prentice- Hall.

Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine.* Cambridge, MA: MIT Press.

Table 1. Illustrates the hierarchy of linguistic functioning in infant language development. Reproduced from Powers (1979)

Age	*Level*	*Level of control*	*Linguistic equivalent*	*Behavioural Output*
24 months +	11	System-concept	Language as a whole; socio-cultural connotations	
18 months	10	Principle	Pragmatics, usage	
12 months	9	Program	Grammar, complete sentences	Two-word utterances emerge
	8	Sequence	Syntactic ordering	Coordination of speech
	7	Category	Naming, semantics	Increased repertoire of words
7-8 months	6	Relationship	Prepositions	
	5	Event	Words, idioms	Coordination of speech mechanisms i.e. jaw, tongue, diaphragm
	4	Transition	Intonation to the ear	
3 months	3	Configuration	Syllable recognition	Babbling
	2	Sensation	Sound quality, pitch	Pleasure vocalisations
	1	Intensity	Loudness	

Figure 1. A negative feedback loop acts continuously to reduce the discrepancy between a reference value/goal and the current perception of disturbance from the environment. Reproduced from Carey (2008) with permission.

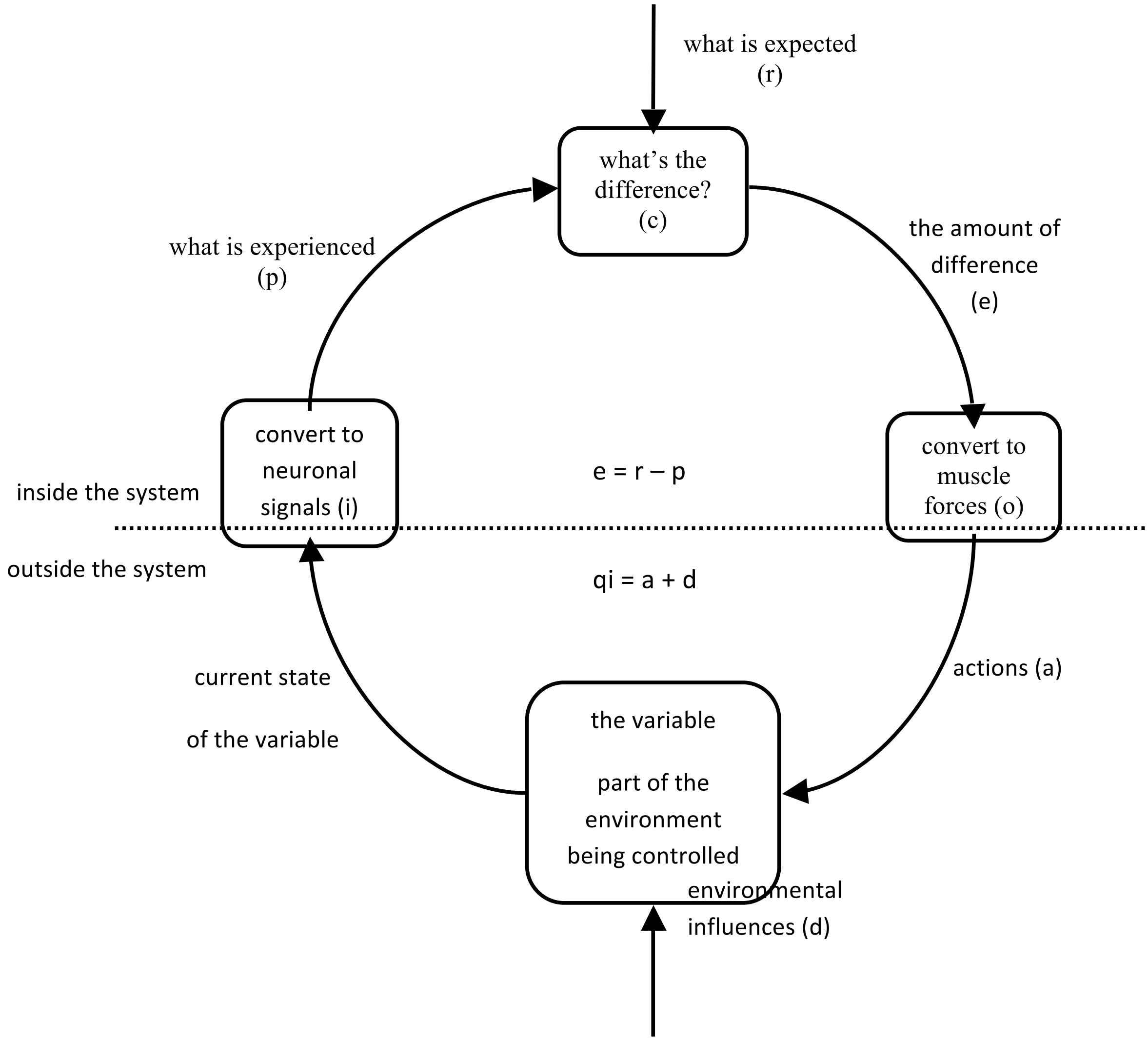

Figure 2. A graph showing the hypothetical stages in developing and reorganising reference frames in infant language development.