DOCTOR OF PHILOSOPHY

Naming and levels of categorisation in young children

Adema, Marleen Th.

Award date:
2008

Awarding institution:
Bangor University

Bangor University, Wales

Prifysgol Bangor, Cymru

Naming and levels of categorisation in young children

Marleen Th. Adema

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology

Bangor University, Wales

April 2008

Acknowledgements

I would like to thank:

My supervisors, Pauline Horne, Carl Hughes, and Fergus Lowe, for guidance, support and humour. I so badly wanted to do a PhD in behaviour analysis. You made it possible, and gave me such a positive environment to work in! Pauline, I was fine, as you predicted. Your endless faith, encouragement, and calmness (!) were crucial.

My friends at Tir Na n'Og Nursery, for support and good times. And to all you lovely children: you made this research such fun! The song says it all.

All the support staff, for help and chats, especially Steve Houlston, for friendly help when the recording equipment turned its back on me.

Vicky Lovett, Sam Hartley, Suzannah Ravenscroft, and Karen Williams for being the best and most committed test assistants I could have wished for.

All the people I met at the conferences before coming to Bangor, for a warm welcome, for ending my isolation as behaviour analyst, and stimulating me to pursue my scientific interests; most notably, Charles Catania (for reinforcing my first attempt to contact other behaviour analysts), Derek Blackman (since that first conference in Amiens you helped me so much to get to Bangor!), and Mecca Chiesa (for sharing your passion for behaviour analysis with me).

Sidney and Janet Bijou, and SABA, for awarding me with a fellowship. Prins Bernhard Cultuurfonds, for funding for my Masters. And the School of Psychology, UWB, for funding my PhD. That all meant so much!

Alle, for always being with me (no matter the distance!), for love, humour, and patience (!), for putting things into perspective, for endless faith in a good end to my projects, and excitement over completed chapters, for giving me the freedom to go and chase my intellectual dreams, and seeing adventure, not disaster, in being apart!

My parents, for love, for always having me in your thoughts, for stimulating me to learn, encouraging me in tough times, and helping me in every way to get where I am now, even though that meant you didn't see much of me over the past years!

Ina, for being the best big sister, moving my stuff from place to place, coming to Wales (from week 1!), and travelling all over the world just to send me postcards!

Nancy, for being a friend and providing a lovely place to be between Bangor and home ever since that first Christmas break.

Vicky, for caring, long chats, emails from Beijing, and the little book of hugs!

Katy, for caring, for always putting a smile on my face, for great times (squash, guitar nights!), and for giving me a Wonderland to escape to!

All my other amazing friends, in The Netherlands (!), and wherever.
I could not have done this without you. You make life fun!

Abstract

Horne and Lowe (1996) suggest that learning a common name for several arbitrary, or disparate, stimuli may establish a category relation between these stimuli. The present studies explore in 3- to 4.5-year-old children whether categories at different levels can be established less directly by training intraverbal relations between lower-level names and potential higher-level category names.

In Study 1a, the children were trained to tact (see animal – say name) eight different "alien" animals randomly allocated to four two-member common name categories. When tested, all 11 children who were tested showed the corresponding untrained listener behaviour (hear /name/ – select animal).

Study 1b linked these lower-level names (hib, feb, tor, and lup) to potential higher-level names (zaag and noom) in an echoic and intraverbal game. Following this, 5 out of 8 children showed correct listener behaviour at the higher name level in Leg 1 of the study, and 3 out of 5 in Leg 2.

In Study 1c, two gestures were trained, one to a zaag, one to a noom. One of the 3 children showed untrained transfer of function to all other aliens, via lower- and higher-level names. Another child showed partial transfer.

In Study 1d, animal cries were trained to two different aliens, again to one zaag and one noom. The one remaining child showed partial transfer.

Study 1e found that a more general verbal prompt than that used in Studies 1c and 1d elicited the appropriate gestures and animal cries, but never both on the same trial.

Studies 1f and 1g found listener behaviour to the gestures and animal cries to be in place.

When tested in Study 1h, the child correctly sorted all stimuli into lower- and higher-level categories.

Studies 2a-h replicated these studies, but employed pre-training with familiar stimuli before the alien Studies 2a, 2b, 2c, and 2h. Alien tact training required fewer trials than in Studies 1a-1h, and 3 more children (7/8, and one more in retests) showed correct listener behaviour at the higher name level in Leg 1, while 3 out of 7 passed for Leg 2. However, transfer of function results were as in Studies 1, while in Study 2h the child sorted stimuli correctly into lower-, but not higher-level categories.

The data are consistent with a Skinnerian account of verbal behaviour.

Contents

Appendices **305**

Chapter 1 – The Development of Verbal Behaviour

> Verbal behavior is undoubtedly the most important and conspicuous behavioural characteristic of humans in relation to other species. From planning a trip to the grocery store, to the development of new cultural practices, to genocide, the influence of verbal relations are seen virtually everywhere in human affairs. It is arguably the most important field for the future of behavior analysis (Leigland, 2007, p. 336).

The study of verbal behaviour is of great importance, also to our understanding of human behaviour in general, as Leigland points out here, and elsewhere (see Leigland, 2001). Greer and Ross (2004) emphasize its relevance to (normal and special) education: "our understanding of verbal behavior is key to the analysis and teaching of new operants" (p. 143). Sundberg and Michael (2001) highlight that verbal behaviour underlies most learning in normally developing children. Therefore, when verbal skills are lacking or defective, the development of these skills is a major goal. Sundberg and Michael describe the benefits of the analysis of verbal behaviour, specifically Skinner's, for children with autism. And Greer and Keohane (2005) give an eloraborate description of how simple and more complex verbal repertoires can be built up by training component skills, thus specifying how the analysis of verbal behaviour can be, and has been, put to use.

Skinner's *Verbal Behavior*

In publishing a 500-page book on verbal behaviour, Skinner (1957) expresses his views on the importance of the topic. In his book *Verbal Behavior*, Skinner

(1957) provides an *interpretation* of many aspects of verbal behaviour in terms of general principles of behaviour. Underlying this undertaking is the assumption that verbal behaviour is continuous with all other operant behaviour[1] (Schlinger, 1995), although Skinner acknowledges that verbal behaviour has "many distinguishing dynamic and topographical properties" (1957, p. 2) and that its study will require a special approach.

It is a causal or functional approach that Skinner presents, rather than a structural one. That is, he is interested in the causes and functions of verbal behaviour, not in the structure of language. He leaves analyses of the latter to linguists. To distinguish his subject matter from that of linguists, he avoids using terms such as "language", "language use", and "language acquisition". The causes of behaviour to be searched for within a functional account should have an acceptable scientific status; we should be able to measure and manipulate them. Therefore, ideas or thought processes should not be invoked as causes. Skinner speaks of "explanatory fictions" (1957, pp. 6-7) in these cases, because these ideas and thought processes need to be explained in turn. Instead, the causes of behaviour should be found in the environment, and in a person's environmental history. Skinner notes that verbal behaviour usually is the result of multiple causes.

Skinner's functional definition of verbal behaviour is "behavior reinforced (or effective) through the mediation of other persons" (1957, p. 2). In line with this definition, he points out that any movement that can affect another organism may be verbal; as examples he lists vocal behaviour, signing, writing, typing, pointing to words, ceremonial language, and even the language of flowers and gems. To refine

---

[1] Operant behaviour is behaviour, which operates upon the environment (Skinner, 1957).

the definition, he later adds: "the 'listener' must be responding in ways which have been conditioned *precisely in order to reinforce the behavior of the speaker*" (1957, p. 225). Skinner acknowledges that although the listener has a role in reinforcing the speaker, the definition does not include listener behaviour as verbal. He confirms that the behaviour of the listener is not necessarily verbal, and states, that "an adequate account of verbal behavior need cover only as much of the behavior of the listener as is needed to explain the behavior of the speaker" (1957, p. 2).

Verbal behaviour should not be confused with its traces, such as speech sounds, transcripts, or texts. Within linguistics it is mostly these traces that are studied (see Julià, 1983). Skinner focusses on *verbal operants* and the variables of which they are a function. Verbal operants are verbal behaviours that have an effect on (or, that operate upon) the environment, which in turn has an effect on the organism. They are *kinds* of behaviour, rather than instances. An example of an instance of behaviour is my neighbour greeting me, yesterday evening at 8 o'clock, while greeting someone, in general, is a kind of behaviour. "Although we observe only instances, we are concerned with laws which specify kinds" (Skinner, 1957, p. 20). A verbal operant, the "unit" of behaviour, does not have a fixed size. The production of a single speech sound can be an operant (e.g., "oh!"), but so can the production of a phrase, or even a whole sentence, as long as the behaviour in question is under the same functional control (that is, a function of the same environmental variables) when it is produced. Every speaker has a verbal repertoire, which is a collection of verbal operants forming the potential behaviour of the speaker. And every operant within that repertoire has a probability of emission within given circumstances. This probability of emission, or strength, of a given operant by a particular person at a given time is the dependent variable within the functional analysis, while the conditions and events in the present and past environment of this person form the independent variables.

When responses are repeatedly followed by reinforcement, this builds and maintains the strength of an operant. This process is called "operant conditioning". Reinforcement can be seen as some sort of reward, which can be tangible or non-tangible. It can range from a biscuit, a stroke on the head or pat on the shoulder, to "well done!", a smile, or even a glance in a child's direction (see also Schlinger, 1995, pp. 179-181). What all these stimuli have in common is that they can increase the strength of operants (see MacCorquodale, 1970, reacting to Chomsky's, 1959, criticism of the term "reinforcement"). And it is only when they are observed to do so, that they can be classified as reinforcers; in other words reinforcers are defined *functionally* with respect to whether they increase the future probability of the responses they follow. Thus, what reinforces one person's behaviour may not have the same effect for someone else, and what reinforces someone's behaviour at one time may not do so at another time. Skinner also notes that automatic reinforcement may occur. He speaks of automatic reinforcement, when a response automatically produces reinforcement. That is, when a child produces a verbal response similar to responses of her caregivers, or simply when she hears the sounds she produces, this in itself reinforces these responses. There is no need then for a caregiver to provide reinforcement.

In *Verbal Behavior*, Skinner (1957) gives a classification of different verbal operants in terms of their functional relation to a stimulus (if present), reinforcement, and other verbal responses. For this classification, he introduces several new terms. The different verbal operants are listed below.

Mand. This is a verbal operant in which the response is reinforced by the consequence it specifies. For example, "help!", "think!", "no more what ifs!" or, "somebody bring me some water", all emitted within certain circumstances. There is no particular relation between the response and a prior stimulus. Instead, the response is controlled by conditions of deprivation, satiation, or aversive stimulation, as is clear in the examples.[2] There are different types of mands, such as commands, requests, advice, and warnings. Mands are established when a child emits a particular verbal response (e.g., "hungry") as a function of deprivation or aversive stimulation conditions, and the verbal response is followed by specific reinforcement, reducing the deprivation or aversive stimulation. Operant conditioning thus brings the reinforcement specifying verbal response under the control of the relevant deprivation and aversive stimulation.

Tact. A tact is a particular response under the control of a specific object or event, or a property of either. An example is when someone sees a CD (for example, having just opened a present), and says "(a) CD!" The presence of the stimulus increases the probability of emission of the tact. A tact is brought about when generalised reinforcement (e.g., "right!") is provided for an appropriate verbal response ("cat") in the presence of a stimulus (a cat). Initially, approximations to the

---

[2] In relation to responses controlled by motivational conditions of deprivation, satiation, or aversive stimulation, Michael and colleagues (see Laraway, Snycerski, Michael, & Poling, 2003) introduce the term "motivating operations" (MOs), previously called "establishing operations" (see Sundberg & Michael, 2001). According to Laraway et al., MOs have two defining effects. "They alter (a) the effectiveness of reinforcers or punishers (the value-altering effect) and (b) the frequency of operant response classes related to those consequences (the behavior altering effect)" (p. 412). For instance, deprivations of liquid would momentarily increase both the effectiveness of water and other drinks as reinforcers, and the frequency of behaviour that, in the past, has resulted in obtaining drink. Skinner discusses conditions of deprivation and satiation in a very similar way, with the same example, in his section on motivation and emotion (1957, p. 31–33, and see Chapter 3). He even speaks of "the operations" we engage in when trying to alter the probability of emission of the response "Water!": "The important events are the operations which are said (by the layman) to change the state of thirst" (1957, p. 32).

appropriate verbal responses (e.g., "at" for "cat"), may be followed by generalised reinforcement. This way, appropriate verbal responses are shaped. Caregivers may also prompt for echoic behaviour (see below) when teaching a child to tact.

Skinner also describes several operants that are under control of verbal stimuli; he distinguishes between echoic, textual and intraverbal behaviour.

Echoic behaviour. We speak of echoic behaviour when a speaker produces sounds similar to a stimulus. For example, the stimulus is /cat/, and the speaker's response is "cat". As Skinner writes, echoics are often taught with help of mands of the type *Say "X"*. The listener (the child) then becomes a speaker, and her verbal response will be reinforced when it has the sound pattern "X". Generalised reinforcement is used to establish an echoic repertoire. Initially, the child's echoic behaviour will be only an approximation to the provided model, but gradually the child's responses will become more similar to these stimuli. Once an echoic repertoire is established, new verbal responses of any type (e.g., mands, tacts, etc.) can be brought about more easily by modelling these responses, and prompting the child to echo. Self-echoic behaviour is also observed in children, which may provide automatic reinforcement (the child hears the sounds she produces, and this reinforces the self-echoic behaviour).

Textual behaviour. A person is said to show textual behaviour when a verbal response is produced under the control of a nonauditory verbal stimulus such as a printed word, or Braille. An example would be when a speaker says "bike" when seeing the letters BIKE. This type of behaviour arises when generalised reinforcement is provided for vocal responses that correspond with stimuli consisting of a letter, or of letter combinations. Caregivers or teachers can also prompt for echoic behaviour in relation to the textual behaviour they model.

Intraverbal behaviour. Intraverbal behaviour is verbal behaviour under the control of other verbal behaviour, without the point-to-point correspondence between stimulus and response that applies in echoic and textual behaviour. An example would be when, in response to /the capital of The Netherlands/, one says "Amsterdam", or in response to /how are you?/, one says "fine thank you". Skinner points out that the alphabet, counting, adding, multiplying and most historical facts are all learned as series of intraverbal responses. Intraverbal operants vary in size. When citing the alphabet, an intraverbal is as small as one speech-sound. When reciting a poem, it can consist of many words. A novel verbal stimulus can evoke one or more intraverbal responses because it resembles other stimuli.

With regard to the establishment of intraverbal behaviour, Skinner makes clear that in a person's lifetime one particular stimulus comes to control many different responses, and one particular response comes to be controlled by many different verbal stimuli. This results from reinforcements "under a great variety of inconsistent and often conflicting contingencies" (Skinner, 1957, p. 74). A verbal stimulus can come to set the occasion for reinforcement of a particular response, for example, when the two verbal forms often occur together. This could be the case because nonverbal circumstances may regularly arise that evoke both forms. The intraverbal behaviour that a person will show depends on the person's verbal history, as does all his or her other verbal behaviour. Therefore, a particular stimulus word can evoke very different responses in different people. Skinner points out that a group of people may show similarities in verbal history, which results for example in most Americans (and inhabitants of other countries with red, white and blue flags) responding to the stimulus /red, white/ by saying: "and blue".

As Catania (1988) points out with regard to all the verbal operants just described, "any utterance, however, is likely to involve these and other relations in combination (…). These elementary relations are only the raw materials from which verbal behavior is constructed" (p. 7). Skinner (1957) notes, that perhaps more than any other behaviour, verbal behaviour is likely to be under the control of multiple causation. He stresses the role of intraverbal, echoic and textual responses, and combinations of these, in sustained speech. And he highlights that "the effect of the speaker's own behaviour leads him to compose and edit what he says and to manipulate it in verbal thinking" (1957, p. 80). In addition to the above relations, Skinner (1957) therefore describes another type of verbal behaviour (another verbal operant).

Autoclitic. Autoclitic behaviour is verbal behaviour based upon, or depending upon, other verbal behaviour. Autoclitic behaviour plays a role in the production of larger verbal responses, such as sentences. Sentences cannot solely be combinations of the elementary units, the basic verbal operants, described above (Catania, 1988). Autoclitic behaviour is involved in the process of self-editing, and in logical verbal behaviour as well. Skinner (1957) identifies two main types of autoclitic: descriptive autoclitics, and relational autoclitics.

Descriptive autoclitics are collateral verbal responses that can give different types of information regarding the speaker's own behaviour, or the controlling relations of initial verbal responses. For example, they can specify which operant they are combined with, whereby "I hear …" is paired with an echoic, and "I ask you …" with a mand. They can also describe different strengths of responses, as in "I guess …" versus "I assure you …". Or they can indicate the speaker's emotional or motivational condition, as in "I am happy to say… ". The verbal community arranges the reinforcement contingencies that bring about descriptive autoclitic behaviour, by asking a child questions such as, "Did you say that?" or, "Why did you say that?" These questions are prompts for self-descriptive behaviour. Other questions, such as, "Did you see it, or did someone tell you?" prompt for comments on the controlling relations, while "Are you sure?" questions the child on the strength of her response, and "Are you happy about that?" prompts for a comment on the child's emotional condition.

Relational autoclitics are verbal responses that modify other verbal responses in a more subtle way than descriptive autoclitics, by playing a part in ordering and grouping of responses. Examples of relational autoclitics are "with", "to", "and", "or", and grammatical tags of inflection such as "-s" in "the boy runs".

One can also speak of relational autoclitic behaviour, when autoclitic "frames" (such as "...and... " , and "the boy's...") are combined with responses that are appropriate to a particular situation. Having learned a few specific instances of verbal responses with regard to two objects ("the cat and the dog", "the mum and the dad", and "the boy and the ball"), the partial frame "...and..." may become available for combination with different responses. Skinner writes that the fully filled out frames are a function of environmental variables, whereby "the relational aspects of the situation strengthen a frame, and specific features of the situation strengthen the responses fitted into it" (1957, p. 336). He calls these autoclitic frames "partially conditioned" (p. 336).[3]

---

[3] Skinner does not elaborate upon this.

Skinner (1957) maintains that verbal operants are *functionally independent*. As he points out, when a child has learned to ask for a doll with the mand "doll!", this does not automatically entail that the child can also tact a doll (1957, p. 187). And if the child does indeed have both the mand and the tact for a doll in her repertoire, these involve different functional relations and each requires a separate account. However, Skinner notes that sometimes when children have learned to tact an object, they may also produce the corresponding mand, apparently without being trained to do so.[4]

Furthermore, different operants, or different mands, cannot be distinguished on the basis of the structural characteristics of the responses observed, as is done in linguistics. They can only be categorised on the basis of the circumstances in which they are emitted, and in the case of mands, on the basis of the behaviour of the listener; if the listener does not provide the reinforcement specified by the speaker, the emitted response did not function as a mand (see Richelle, 1976).

Sometimes it is hard to determine what type of operant is observed, because, especially in the case of mands and tacts, there can be a mixture of controlling relations. For example, when someone says, "oh, we're out of milk!" this is a tact of the situation in the fridge. But when, in reaction to this remark, someone else goes out to buy some milk, it has also served as a mand. Skinner (1957, p. 151) calls this an *impure tact*. Similarly, when mother and child are seated at the breakfast table and mother asks the child, "What would you like?" the child may reply by saying, "Chocolate spread". If mother then puts chocolate spread on the child's toast, the child's verbal response has served as a mand. But because the child said "Chocolate

---

[4] Horne and Lowe (1996, p. 211) write that, "once naming of an object is established, then, after direct training on only a few mand relations, name-manding often follows".

spread" in the presence of chocolate spread, controlling relations of a tact are also involved. Therefore, one could speak of an impure mand here (see Oah & Dickinson, 1989; and see Sundberg & Michael, 2001, pp. 710-711, who speak of a "pure mand" when a requested item is not present, as opposed to a functional relation that is part tact and part mand).

Chomsky's review

Skinner's book *Verbal Behavior* received many favourable reactions (see, for example, Dulany, 1959; Morris, 1958; and Osgood, 1958. See also Knapp, 1992). And then, there was Chomsky's review (1959).

In his lengthy review, Chomsky heavily criticized Skinner's book. His three main points were as follows (see also Chomsky, 1965):

1) The behavioural approach is too simplistic to explain the complexity of language learning.
2) The behavioural approach cannot explain the generativity that is observed in children who are learning language. Direct training and reinforcement is not necessary for every single item in a child's repertoire (see Schlinger, 1995).
3) The poverty of the stimulus is a problem; the language environment of the developing child is chaotic, and therefore an innate language acquisition device (LAD) is needed (see Dale, 2004).

These criticisms[5] led to the exploration of new directions within behaviour analysis and the field of verbal behaviour.

Verbal behaviour: new directions

Three new approaches were developed, building on Skinner's *Verbal Behaviour*: stimulus equivalence theory, relational frame theory, and the naming account. All three are behavioural accounts of categorisation.

Categorisation can be defined as responding in a similar manner to different stimuli, whereby the stimuli can have similar features (as is the case for the category of cats), or show no or hardly any physical similarities (e.g., the category of furniture). As Lakoff notes, "Categorization is not a matter to be taken lightly. There is nothing more basic than categorization to our thought, perception, action, and speech" (1987, p. 5).

Categorisation is closely linked to the issue of generativity, raised by Chomsky (see above). We can respond appropriately to new stimuli without being trained to do so, because these stimuli belong to the same stimulus class as previously encountered stimuli (having similar physical features, and/or being members of a common name relation), and thus the new stimuli occasion behaviour learned with regard to the previously encountered stimuli.[6] For example, when a

---

[5] See MacCorquodale (1969, 1970), and Richelle (1976, 1993) for reactions to Chomsky's review. See Palmer (2000a, 2000b), and Schoneberger (2000) for extensive behaviour analytic reviews of Chomsky's nativist approach. And see Andresen (1990, 1992), Czubaroff (1988, 1989), Lee (1984), Richelle (1993), and Stemmer (1990) on the discussion between Chomsky and Skinner.

[6] It is tempting to say that we respond similarly to different stimuli *because* we have categorised them. But that would result in a circular formulation, because we have just defined categorisation as responding similarly to different stimuli.

child first sees a snake, a caregiver will tell her how to respond to it. One particular snake belongs to the stimulus class of snakes. We learn features of classes of stimuli, and that prepares us for situations in the future containing similar stimuli. Why this is useful, is especially clear in this example of snakes. During the course of our lives we regularly come across similar situations, and often our behaviour is shaped further and further to become ever more appropriate towards specific classes (or subclasses) of stimuli.

Categorisation is a topic that is widely studied within developmental psychology (see, for example, Rakison & Oakes, 2003)[7]. Stimulus equivalence theory, relational frame theory, and the naming account attempt to analyse categorisation and describe how it is established, from a behaviour analytic perspective.

The naming account, with its extensive description of how verbal behaviour may develop over the first two years of life, is the main focus of this thesis, and is therefore described in detail below. However, first stimulus equivalence theory and relational frame theory are described to provide the theoretical context in which the naming account was developed.

Stimulus equivalence theory

In the 1970's, Sidman and colleagues developed stimulus equivalence theory. Sidman (1971) put forward the term 'stimulus equivalence' in an attempt to explain the finding that in certain match-to-sample tasks people seemed to relate presented

---

[7] And see Bruner, Goodnow and Austin (1956), and Neisser (1987) for remarkably behavioural definitions of categorisation (coming from cognitive psychologists), in line with the definition given here.

stimuli in a bi-directional, symbolic way[8], though they were not taught to do so (see Sidman, 1992, and 1994, for an introduction to this account).

Match-to-sample tasks are procedures in which a participant is required to match a sample stimulus to a particular one of the provided comparison stimuli. The number of comparison stimuli can vary across experiments. An example of a match-to-sample trial is when a participant is presented with the spoken word "car" as a sample stimulus, and pictures of a car, a key, a dog, and the sun, or the printed words CAR, KEY, DOG, SUN. Typically, match-to-sample studies start with baseline training. For example, the participant is presented with the spoken word "car" (stimulus A), and is trained to select the picture of the car (stimulus B) in the first combination of comparison stimuli, and the printed word CAR (stimulus C) in the second combination. In short, the relations AB and AC are trained. Following baseline training, participants are tested. In tests, the elements within these relations can be reversed, so that a sample stimulus becomes a comparison stimulus and vice versa (BA, and CA), and elements from different relations can be combined (BC).

Sidman described three emergent relations observed in these match-to-sample studies: reflexivity, symmetry and transitivity. He characterised the relations in mathematical terms. *Reflexivity* (or identity matching) is observed when participants, in the absence of direct training, match a stimulus on to an identical stimulus (i.e., if A then A, if B then B). *Symmetry* is observed when after being trained on the conditional relation 'if A then B', people demonstrate that the relation also holds when the elements are reversed, that is, when presented with sample stimulus B they select A (if B then A). And, *transitivity* is observed when after training on the two

---

[8] There is a 'symbolic' relation between two stimuli, when one stimulus *refers to*, or *stands for* another (see Horne & Lowe, 1996, p. 201), as is the case, for example, for the printed word CAR or the spoken word "car" in relation to a picture of a car.

conditional relations ‘if A then B’ and ‘if B then C’, participants presented with A select C (‘if A then C’). These three relations are called ‘emergent’ relations, because they were found to develop without direct training within the context of the experiment. The combination of these three emergent relations shows that equivalence is established between the stimuli A, B, and C. The stimuli form an equivalence class, and the trained and tested relations are equivalence relations.

The stimulus equivalence research was seen as opening up new opportunities for a return of behaviour analysis to the research area of symbolic categorisation and, more generally, verbal behaviour. This was because stimulus equivalence research seemed to complement the behaviour analytic account of verbal behaviour on just those issues that, according to the critics, it was insufficient (as discussed above).

A problem was, however, that the emergent relations could not readily be explained in terms of basic learning principles. Therefore, an alternative explanation needed to be found. Sidman suggests that equivalence may be an “unanalysable primitive in the description of behaviour” (1990, p. 111), as is the case for reinforcement. Primitive stimulus functions are not derivable from other behavioural processes (Sidman, 1992). In other words, equivalence relations may be a biological “given”; they emerge as a result of specific reinforcement contingencies (Sidman, 2000), such as those provided in match-to-sample training, “because contingencies of survival have made us that way” (1990, p. 113).

Sidman (2000) explains that the reinforcement contingencies (arranged in baseline training of match-to-sample studies) produce analytic units and equivalence relations. The analytic, or analysable, units can consist of two, three, four, five, or more terms. In *operant reinforcement*, there is a two-term contingency, in which a defined response produces a defined reinforcer. In *simple discrimination* there is an additional third term, the discriminative stimulus. Only in the presence of this stimulus will the response produce the reinforcer. In *conditional discrimination* there is an additional fourth term, a conditional stimulus. In the presence of the discriminative stimulus the response will produce the reinforcer, but only if a specific conditional stimulus is present. A contingency can have even more terms, for example in the case of second-order conditional discrimination. As noted, according to Sidman, reinforcement contingencies also have equivalence relations as outcome, "these consist of ordered pairs of all positive elements that participate in the contingency" (2000, p. 128). All the elements of the analytic unit or contingency (stimuli, responses, and reinforcers) enter into the equivalence relations, and form equivalence classes. Such potential relations arise when there is no conflict between the contingencies and the equivalence relations.

There is a complication when several different simple or more complex relations (e.g., the conditional relations A1B1, and A2B2) are trained with a common reinforcer and/or a defined response (e.g., key pressing), which is possible. This would result in all stimuli involved in these relations becoming part of one large equivalence class (e.g., A1B1A2B2). Sidman (2000) suggests that this is initially what happens, but that the contingencies then demand that the common response and reinforcer elements selectively drop out of the class, so that the analytic unit (the *n*-term contingency) can be formed. He notes that this problem can be avoided by using different responses and different reinforcers to establish separate relations. Sidman speaks in this case of "contingency-specific reinforcers and responses" (2000, p. 137).

As Sidman (1994, and 2000) points out, within equivalence classes the distinction between stimuli and responses is removed.

Sidman (2000) also reacts to the question whether naming is a critical determiner of the emergent performances that define equivalence relations (as suggested by Dugdale & Lowe, 1990; and Horne & Lowe, 1996):

> Any name we apply to stimuli is a defined discriminative response. Our theory states explicitly that any defined response components of the contingencies have a status that is equal in every way to the stimulus and reinforcer members of the classes. Although just as important, responses require no separate treatment. I think this provides a simple but satisfactory resolution to the naming controversy (2000, p. 145).

So names, or verbal rules for that matter, can enter into the equivalence relations as responses but should not be assigned any special status, or a special mediating function (see Sidman, 1994). Sidman (1990) notes that when participants in match-to-sample tests use names or verbal rules without being taught to do so, the question arises what came first. It may well be that equivalence gave rise to the names or verbal rules, rather than the other way around. If, however, it was the names or rules giving rise to equivalence, then it remains to be explained where they came from. Sidman (1992) adds: “That rules can give rise to equivalence relations does not mean that equivalence relations require rules” (p. 21). Stimulus equivalence can indeed be established by verbal rules or names, but it can also come about through experience with contingencies (Sidman, 1992). All in all, one main point of Sidman’s (1994) account is that verbal skills are not necessary for stimulus equivalence. Related to this point, is that Sidman thinks that stimulus equivalence may also occur in other animal species, and that if it were demonstrated in nonverbal animals, this would rule out the primacy of verbal behaviour in the development of stimulus equivalence. So far, stimulus equivalence has not been demonstrated unequivocally in animals (but see Schusterman & Kastak, 1993; and Horne & Lowe, 1996, p. 223-224; Horne & Lowe, 1997).

Relational frame theory (RFT)

Two behaviour analysts who did not agree with Sidman and his colleagues, are Hayes and Hayes (1989, 1992), later supported by Barnes-Holmes and colleagues (e.g., Barnes-Holmes, Barnes-Holmes, Smeets, Cullinan & Leader, 2004). Hayes and Hayes developed relational frame theory (RFT), which states that stimulus equivalence and other verbal activities, such as naming, understanding, analogy, metaphor and rule-following, are the result of a learning process of relational responding (Barnes-Holmes, et al., 2004; and Hayes & Berens, 2004).

A *relational frame* is defined as "a specific class of arbitrarily applicable relational responding" (Hayes, Barnes-Holmes & Roche, 2001, p. 33)[9], and this relational responding is derived, controlled by context, and the result of a learning history (Hayes, 1994; Hayes, Barnes-Holmes & Roche, 2001).

A relation between stimuli is an *arbitrarily applicable relation* when one responds to the relation between the stimuli on the basis of contextual cues, rather than the physical form of these stimuli. To apply an example of Hayes (1991) to the British coin system: a young child will work harder for a 10 pence coin than for a 20 pence coin, because the 10 pence coin is physically larger. The child's relational

---

[9] Or, as Palmer (2004, p. 193) puts it: "a relational frame (...) is a set of responses that relate classes of stimuli".

responding to the two coins is non-arbitrary, purely based on the physical form of the stimuli (i.e., their size). An older child will work harder for the 20 pence coin, thereby responding to the convention she has learned that although this coin is physically smaller, it is worth more. Provided with the two coins as stimuli, and the contextual cue of the word "more", she would select the 20 pence coin. In this case, the child's relational responding to the two coins is arbitrary relational responding. As this example makes clear, children learn to respond to stimuli in relation to other stimuli and contextual cues.

Hayes (1994) describes three defining characteristics of relational frames (arbitrarily applicable relations)[10]. These characteristics are:

(1) Mutual entailment. A particular relation between stimulus A and B implies a reciprocal relation between B and A. For example, if "A is more than (or the same as, or opposite to) B", then "B is less than (or the same as, or opposite to) A". Thus the relation is mutually entailed, and these relations are controlled by contextual stimuli. The relation of sameness is comparable to symmetry as defined by Sidman.

(2) Combinatorial entailment. A particular relation between stimuli A-B and B-C implies not just a relation between B-A and C-B (by mutual entailment), but also a relation between stimuli A-C and C-A. For example, if "A is bigger than B" and "B is bigger than C", then "A is bigger than C" and "C is smaller than A". Combinatorial entailment is comparable to transitivity as defined by Sidman.

---

[10] Blackledge (2003) notes that although the term "relational frame" might suggest a structure, it is better to think in terms of a process, relational framing: "People frame events relationally in the moment as an active process that is a function of their extensive learning history and stimulation in the present environment. 'Storage' of these frames as structures is not implied and not required" (p. 429).

(3) Transformation of stimulus function. When stimuli A and B are involved in a relational frame, functions of one of these stimuli may transfer via the relational frame to the other stimulus involved (Barnes-Holmes et al., 2004). For example, if someone is offered a glass of apple cider but turns down the offer because she doesn't like alcohol, a friend may say: it's just like apple juice. By specifying a relation of sameness between apple cider and apple juice, the functions of apple juice (e.g., the girl in the example would drink it when offered) can transfer to apple cider. When more than one stimulus is involved in the frame (for instance, also a stimulus C), the function of one stimulus may transfer to any or all of the other stimuli.

As the examples above suggest, there are several different types of relational frames such as the frames of coordination, opposition, distinction, comparison, hierarchical relations, temporal and spatial relations, and deictic relations. (For a description of a typical relational frame experiment, see Palmer, 2004, pp. 191-192.).

Hayes, Barnes-Holmes, and Roche (2001, pp. 43-44) also offer a new definition of verbal behaviour in terms of relational frames: "Verbal behavior is the action of framing events relationally", and they add, "verbal stimuli are stimuli that have their effects because they participate in relational frames".

Hayes (1994) suggests that stimulus equivalence is the result of the application of one particular type of relational frame, that of "co-ordination". In this frame, stimuli are related on the basis of sameness (that is, similarity), whereby sameness is determined by contextual cues, rather than the stimuli having the same physical form (as in the example of the coins above). As for the role of naming in the establishment of stimulus equivalence, proponents of RFT pose that the core-defining element in both stimulus equivalence and naming is relational responding (Barnes-Holmes et al., 2004). On the other hand, Hayes (1994) points out that naming and verbal rules can help establish stimulus equivalence, in the sense that names can be contextual cues for relational responses. However, Hayes does not consider naming to be necessary for stimulus equivalence; equivalence can also be generated directly by contingencies, without there being rules or names involved. Furthermore, according to Hayes and Hayes (1989), listener behaviour training in the absence of naming (so without training speaker behaviour) should also help in bringing about relational responding such as equivalence (see also Luciano, Gómez Becerra & Rodríguez Valverde, 2007).

So how does arbitrarily applicable relational responding (relational framing), and the resulting stimulus equivalence, come about? According to RFT it comes about through *multiple exemplar training*:

> (...) it seems that relating as an overarching class could be formed in a way somewhat similar to that of generalized imitation—through exposure to multiple exemplars across a variety of situational contexts that refine the nature of the response and sources of stimulus control over it (Hayes et al., 2001, p. 25).

RFT was criticised for lacking a *detailed* description of the learning process through which stimulus equivalence and relational frames are established (see Horne & Lowe, 1996). Palmer (2004) shares this criticism. In his review of Hayes et al.'s (2001) book, he asks with regard to multiple exemplar training, "but what counts as an 'exemplar'?" And he comments:

> (...) it is not enough to refer to a history of multiple exemplar training, for that expression embraces a wealth of mysteries. (...) much more detail is required about the variability of the history and its relation to variability in behavior. In

> the case of relational frames, what are the critical features of the relevant history, and precisely what behavioral effects follow? (2004, pp. 195-196).

Hayes and Barnes-Holmes' (2004) response to this comment is that it is included in the definition of operant response classes that they result from contingent reinforcement (the authors refer to Catania, 1998). The authors maintain that it is completely in line with this, to say that for a relational operant the relevant relational properties are brought about through a history of reinforcement. Hayes and Barnes-Holmes admit that,

> (...) it is fair to point out that the specific histories that give rise to relational frames have not yet been fully demonstrated, but that is not because those histories are mysterious: it is because the relevant demonstrations are an empirical matter (...) But conducting such empirical work requires considerable clarity about the unit of behavior that is being shaped—clarity about that unit and its implications are precisely why Relational Frame Theory is needed at this point (2004, pp. 216-217).

Recently, Luciano et al. (2007) published a paper on the effects of training multiple exemplars of object-word and word-object relations on generalized listener behaviour and the establishment of derived equivalence. The authors suggest that their results support relational frame theory, and disprove the naming account. Because of the relevance of this paper to both accounts, and the strong claims, it will be discussed here in more detail.

Luciano et al. report three experiments with a girl, Gloria, who was nearly 16 months old at the start of the experimental work, and nearly 24 months old by the end of it.

First, a pre-test (Test 1) was administered in which three objects with which the child was unfamiliar were each presented three times. In the three trials for Object 1, the vocal name was presented, while for Object 2 a gesture (a nonvocal name) was presented. Thirty minutes later the child was asked to give the name for each of these objects. Following the object-name presentations for Object 3, the child was prompted for a "receptive symmetrical response" (listener behaviour to the name); that is, she was asked to select the correct one from three objects upon hearing the name ("give me …"). In none of the three test trials did the child produce a correct response. The authors conclude (in their abstract) that Gloria showed no evidence of a receptive symmetry repertoire or a naming repertoire.

In Experiment 1, multiple exemplar training was given in receptive symmetrical responding, or listener behaviour (both immediate and delayed); this training lasted for one month. Ten novel objects were employed. In training trials, the mother presented the child with an object and its name (see object and hear name), followed by an immediate listener test (hear name – select object), and repeat listener tests at delays ranging from 3 min. to 12hr. In the listener test trials, the target object was presented simultaneously with one previously used object and one object that the child had not yet seen in the experimental context. After reaching the criterion of 7 consecutive correct trials with at least 6 of the 10 different training objects, the child was tested by the mother with 6 new objects (Test 2). Each new see object and hear name relation was presented once, and the corresponding name-object (listener) relation was tested three hours later, again with three objects to choose from on each listener test trial. In Test 2, Gloria (nearly 17 months old now) showed generalised delayed receptive symmetrical responding (i.e., listener responding following only exposure to the relevant object-name relation) for five out of six new objects, but in a naming test with two additional novel stimuli, no vocal or gestural naming, or approximations to either were observed.

In Experiment 2, which began when Gloria was 17 months old, four visual-visual conditional discriminations (A1-B1, B1-C1, A2-B2, and B2-C2) were presented in a match-to-sample format with two comparison stimuli. The match-to-sample procedures employed yet further stimuli, which it is claimed were unfamiliar to the child. In the first two training trials for each relation, the sample stimulus was presented together with the correct comparison stimulus only. After that, a second comparison stimulus was introduced. In correction trials, the sample stimulus was also presented with the correct comparison only. In training and test trials the sample stimulus was held up, while the comparison stimuli were presented in a box. Training took about 2 months. When tested, the child showed derived visual-visual symmetry (BA and CB relations emerged) and transitivity/equivalence responding (AC and CA relations emerged). At the end of the experiment, when Gloria was about 19 months old, she was also tested again for naming and listener behaviour ("emergent symmetry") with regard to two novel objects (Test 3) following exposure to the two relevant object-name relations. Her responses were incorrect on the four naming trials, but correct on the two symmetry trials.

Experiment 3 started when Gloria was 22 months old. The procedure was mostly as in Experiment 2 except that there were three rather than two comparison stimuli, and new stimuli were used. The child showed the same derived listener relations after training as she did in Experiment 2, and was then tested one last time for naming of two new objects (Test 4), at the end of Experiment 3 (she was now nearly 24 months old). For the first time, she named these stimuli correctly on all four trials.

Throughout all three experiments, no imitative responses or naming responses were observed during training or testing.

In relation to Experiment 1, Luciano et al. note that this seems to be the first demonstration of generalized symmetry in an infant established through multiple exemplar training, in the absence of a spontaneous imitative and naming repertoire.

Furthermore, the authors state that "the results of the second experiment meet the conditions proposed by Horne and Lowe (2000) as necessary to disprove their naming account, that is, the emergence of equivalence responding after listener training, without a speaker component or naming repertoire" (2007, p. 363). In short, Luciano et al. claim that the results of Experiment 2 show that equivalence is possible without naming, and they base this claim on Test 2 and Test 3 in which the child was tested for naming new objects and exposure to the relevant object-name relations.

Finally, the authors admit that in Experiment 3 naming may have played a role, because by the end of that study the child was able to name other new objects in the delayed trials of Test 4, after only one presentation of object and object name.

Several issues can be raised when evaluating Luciano et al.'s results. The results they have obtained with only 1 child require replication, as the authors themselves agree. Also, the number of trials per test is extremely small, and in Experiment 2 and 3 it varies per test, and even per relation.

Another matter, also addressed by the authors, is that the mother conducted all training and test sessions, and what is more, the mother and the father were responsible for coding (analysis and scoring) of all sessions. It would have been better if sessions had been conducted by an experimenter, and coded by at least one independent observer (who, in addition, was blind to the purpose of the study, as

Luciano et al. say the parents were). This could be done on the basis of the recordings of the sessions.

Even more serious an issue is that the child's verbal skills at the start of the study are described solely on the basis of the parents' informal observations. The parents were given guidelines on what to observe in the child's behaviour, but no independent observations were made, nor were any objective language assessment tests done, at any point. Thereby, the authors leave a question unanswered that is perhaps the most crucial question in this study, both methodologically and with respect to the evidence against the naming account and its predictions. This question is, how well listener behaviour, echoic behaviour and naming were developed at the start of the study, and how these behaviours developed during the study. The informal observations by the parents suggest that all three behaviours were shown already at the start of the study, although echoic behaviour was infrequent and usually under instructional control, and the child was only able to name a few common objects and people, such as her 'mum' and 'dad', the family dog, and a teddy bear. The authors suggest that "Gloria's vocabulary of a few words for everyday objects and persons (mum, dad, etc.) is clearly not sufficient to qualify as naming" (Luciano et al., 2007, p. 362). However, if the child consistently produces the name in the presence of these objects and persons when the name is not provided as a verbal stimulus, this would indeed fit the definition of naming.[11] It is important to note as well, that over the course of the eight months in which the three studies were conducted, this naming repertoire will have developed enormously, so much so

---

[11] It fits the definition of the tact (Skinner, 1957), as described earlier. And in Horne and Lowe's (1996) naming account, the tact completes the naming circle when listener behaviour and echoic behaviour have already been established; when all three behaviours are in place with respect to particular objects, the child can be said to name these objects. The naming account is described in detail in the next section.

that by the end of the third experiment the child was able to name an object in a delayed test, after only hearing the name once in the presence of the object.

The ability to name new objects as measured (without positive results) in Tests 1, 2 and 3, is not what is most relevant to the child's performance in the equivalence test trials, and in relation to the naming account. What is relevant is whether the child was able to name the objects used in these equivalence test trials. Luciano et al. described these objects as new, and unfamiliar to the child. However, among the objects used in training and testing in Experiment 2 are, for example, a bunny, a teddy bear, a dog and a construction block. It is not at all unlikely that the child was able to name some, if not all, of these toy objects that feature in most young children's play environments. In fact, the parents reported that the child was able to name her teddy and the family dog, already at the start of the research. In Experiment 3, it is also possible that the child was able to name the starfish (e.g., as "star"), the flower, the pig, and the box. In general, if the child was not able to name the objects, she may very well have been able to name features of the objects. Training of stimulus relations took between 32 and 91 trials in Experiment 2, and between 12 and 54 trials in Experiment 3, so there were many opportunities for the child to develop names for the objects used in training and testing. Even if the child did not overtly name objects or their features, it is still possible that the child named these covertly. It seems a fundamental flaw in the study that no naming tests were done for these objects at the end of each experiment.

In addition, a few comments should be made on Experiment 1. In effect, the receptive symmetry training in this experiment is listener behaviour training (as the authors themselves explain), conducted in much the same way as in studies reported by Horne, Lowe and Randle (2004), and Horne, Hughes and Lowe (2006). The child is trained and tested on listener behaviour, and tested on generalized listener behaviour with new objects, which the child does indeed show.

It should be noted that nothing in this contradicts the naming account, which proposes, as will be seen in the next section, that children show listener behaviour before they can name objects, and that they will become more and more skilled at this. It is not at all unthinkable that at some point, after extensive experience in learning listener behaviour through multiple exemplar training, it may be enough for them to see an object and hear its name produced in the presence of the object once or twice, to be able to display appropriate listener behaviour, as Luciano et al. show. The naming account does not state that naming is necessary to learn listener behaviour.

In light of these issues, the firm statement regarding having met the conditions necessary to disprove the naming account seems slightly premature, and coupled with the fact that the experiment was conducted with only one participant, at this stage the findings should be treated with caution.

The naming account

The naming theory by Horne and Lowe (1996, 1997; Lowe & Horne, 1996;) is another behaviour analytic approach to categorisation, critical of both Sidman's and Hayes and Hayes' approach. Like the other two approaches, it builds on Skinner's (1957) book *Verbal Behavior*. Making use of basic behavioural principles, Horne and Lowe (1996) describe how (aspects of) verbal behaviour can be learned. Within their account of the development of naming and classifying behaviour there is an important role for listener behaviour, and reinforcement is *not* necessary for every single instance of behaviour.

Horne and Lowe see the name relation as the basic unit of verbal behaviour. Naming can be seen as a fusion of listener behaviour, echoic behaviour and tacting. And one can only speak of *naming* when all three behavioural relations are in place. A more detailed definition of naming is given later. First, Horne and Lowe's account of the development of naming will be summarised.

Based on a wide range of developmental and behaviour analytic research, Horne and Lowe (1996) describe how naming can develop from earlier, prelinguistic behaviour. The child first needs to acquire many component skills, such as discriminating between speech sounds and other sounds, discriminating between speech sounds of her own language and other languages, and discriminating between the individual parts of her language (i.e., words). The verbal community often adapts its verbal behaviour, thus making it easier for children to learn the above discriminations; for example, by speaking slowly, using a simplified syntax, and repetition (see, for example, Snow, 1977; Snow & Ferguson, 1977; and Fernald, 1992). The component skills are necessary for the child to learn listener behaviour, which in turn is crucial for the development of naming.

Listener behaviour. Once the component skills are in place, listener behaviour comes about when the child repeatedly hears a particular vocal stimulus, an object name, in the presence of a particular object. When caregivers use social reinforcement to teach the child conventional behaviour with regard to the object, the vocal stimulus eventually becomes a discriminative stimulus for the child to engage in this conventional behaviour in the presence of the object. The caregivers facilitate this learning process by tuning into the behaviour of the child; for example, by naming the object the child is looking at, or by indicating which object or event they name; that is, by pointing to it, or looking at it. Infants learn to follow the caregivers' gaze and pointing (e.g., Morales, Mundy & Rojas, 1998), and later learn to point to objects and events themselves, as well as to pick up objects, to which caregivers often respond by providing the relevant name. In addition, caregivers show infants the conventional behaviours towards objects, and will reinforce the infant's attempts to imitate these. Gradually, the verbal community shapes conventional listener behaviour. The child learns to react in different ways to different objects and events; when seeing a ball, she may throw it, seeing a rattle, she may shake it, and seeing a shoe, she may put it on her foot. And conventional listener behaviour, such as orienting to, pointing to and picking up particular objects (e.g., a shoe), comes under control of specific vocal stimuli ("where's the shoe?"), when caregivers provide these vocal stimuli, model the appropriate behaviour in the presence of the object, and reinforce the child's imitations of it. While initially the child may just imitate the caregivers' behaviour, eventually the vocal stimulus itself is enough for the child to reliably show the appropriate behaviour. At that point, the child has learned listener behaviour.

Listener behaviour may at first be shown only in particular contexts (e.g., when mother holds up the child's shoe), with one specific exemplar of an object (e.g., the child's shoe), and in response to a vocal stimulus ("where's the shoe?") uttered only by the child's caregivers, but later it will generalize to other contexts, other exemplars, and to vocal stimuli uttered by any speaker. When the child learns to generalize listener behaviour to other exemplars (other people's shoes), she starts responding to stimulus classes. Caregivers correct, or simply do not reinforce, in the case of underextensions (that is, when the child does not respond appropriately to different kinds of shoes) and overextensions of stimulus classes (e.g., when the child generalizes the listener behaviour also to socks). In the same way as listener behaviour to the vocal stimulus "where's the shoe" is established, the listener repertoire is further extended to include appropriate reactions to different types of vocal stimuli, such as "pick up ...", or "give me ...", combined with different object words, such as "car", "cup" and "spoon".

At this early stage of development the child's listener behaviour is controlled solely by other people's speaker behaviour.

Echoic behaviour. At around 6 months, infants start babbling. Research with infants who have hearing difficulties suggests that hearing speech sounds in the environment influences the onset as well as the form of babbling (see, for example, Oller & Eilers, 1988). Gradually, infants get more control over their speech musculature, and learn to produce closer approximations to verbal stimuli. Caregivers prompt infants to produce echoic behaviour by saying, "say shoe" and they reinforce the infant's approximations to the verbal stimuli they provide, thus shaping the infant's echoic behaviour, and increasing the probability of vocal imitation in general.

Once caregivers have reinforced several echoic relations, there is some evidence that generalized echoic behaviour develops (see Poulson & Kymissis, 1988; Poulson, Kymissis, Reeve, Andreatos & Reeve, 1991). That is, the infant may echo verbal stimuli without the need for reinforcement from caregivers. Skinner (1957) points out that there may be automatic reinforcement when self-produced utterances match verbal stimuli heard in the environment. And Palmer highlights that it is because of this reinforcing effect of *achieving parity* with the practices of the verbal community that also "one's own utterances can shape and maintain one's behavior" (1996, p. 289). The child may echo verbal stimuli produced by others, but may show self-echoic behaviour as well. And echoic behaviour may occur overtly, or covertly; that is, it may not be perceivable by others.

As Horne and Lowe (1996) describe, when the child learns echoic behaviour, an interaction develops between the child's own speaker and listener behavior; when she produces echoic behaviour, she will hear her own utterance, and in reaction to this verbal stimulus she may then show the appropriate listener behaviour (orienting to, pointing at, or picking up) which she had already learned. The child has now become, as Horne and Lowe put it, "a speaker-listener with respect to her own verbal stimuli" (1996, p. 197).

Naming. Echoic behaviour often occurs in the presence of objects. For example, caregivers prompt a child to say "shoe", when a shoe is present. And while prompting, they may point at the shoe, or their prompt may be a reaction to the child pointing at it (e.g., Tomasello & Todd, 1983). When the child repeatedly sees a shoe while hearing /shoe/, and saying (i.e., echoing) "shoe"[12], this sets the occasion for the development of a third functional relation, in which objects gain control over the child's verbal behaviour. This relation is the *tact*. Skinner (1957, p. 81–82) defines the tact as a verbal response of a given form, evoked or strengthened by a particular object or event, or by a property of either. When the child consistently says "shoe" when seeing her shoe, without anyone providing the verbal stimulus /shoe/, we can say that the child has learned to tact her shoe. The shoe has become a discriminative stimulus for her saying "shoe". And when she also produces this verbal response in

---

[12] Both hearing the verbal stimulus /shoe/ from a caregiver, and hearing her own utterance, occasion listener behaviour (e.g., looking at the shoe).

the presence of other shoes, she has learned to tact shoes in general—that is, shoes as a stimulus class.

Horne and Lowe (1996) add that the child has now learned to *name* the shoe, or shoes as a stimulus class. The tact completes the name relation, which can be drawn as a circular relation, incorporating listener behaviour, echoic behaviour and tacting (see Figure 1.1). The relation is "circular", in the sense that hearing /shoe/ can make the child look at the object, and seeing the object can in turn evoke saying "shoe" which makes the child hear /shoe/ again, et cetera. As Horne and Lowe summarize: "naming involves the establishment of bidirectional or closed loop relations between a class of objects and events and the speaker-listener behavior they occasion" (1996, p. 200).

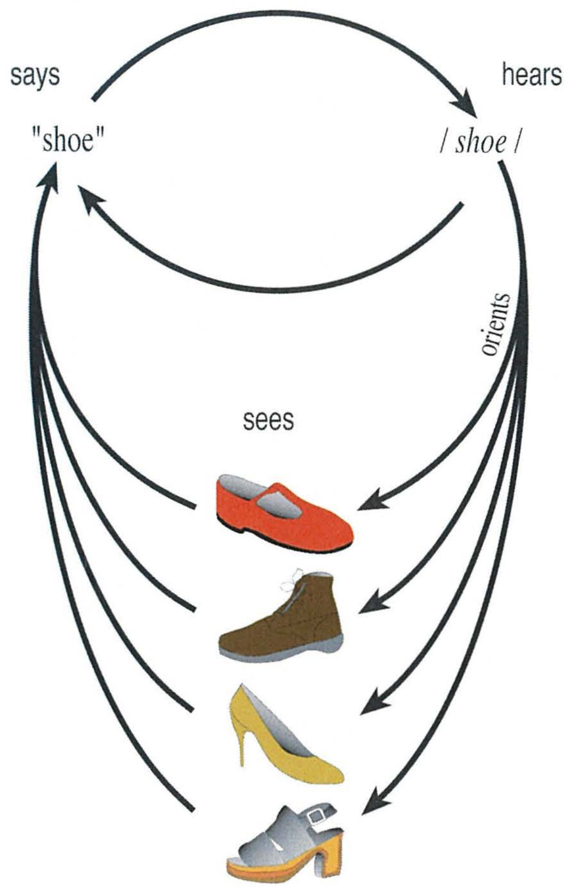


Figure 1.1 (reproduced with the authors' permission, from Horne & Lowe, 1996, p. 201). When a child who has learned to name shoes, sees a shoe, she says "shoe". Upon hearing this self-produced verbal stimulus /shoe/, she shows listener behaviour; she orients, not just to one shoe, but to any of the shoes in her environment that are part of her existing listener behaviour class. And either the seeing of these shoes, or her hearing her own verbal stimulus may then again evoke her verbal response "shoe". In sum, naming can be evoked by seeing a shoe, or hearing /shoe/, and can be re-evoked by seeing a shoe again or through self-echoing. This illustrates the bidirectional relations between a class of objects and the speaker-listener behaviours they evoke.

As seen above, there is an interaction between the three component behaviours—tacting, echoing, and listener behavior—that together constitute naming. The authors emphasise that the properties of naming, and the effects of naming on other behaviour, go beyond those of each of the component behaviours. They point out, for example, that the tact relation between a stimulus and a response is unidirectional and nonsymbolic. Horne and Lowe (1996, see p. 201, and 213–215) define "symbolic" as embodying reference, representation (the name "*re*-presents" the object) and meaning. The tact relation lacks the defining characteristics of symbolic behaviour in that it does not *represent* a stimulus, and it does not *refer* to it or *mean* it. Unlike the tact relation, the name relation does have these characteristics of symbolic behaviour.

Furthermore, the interaction between the three component behaviours that together comprise naming has an important role in the development of the child's naming repertoire. For example, after having learned listener behaviour towards several different types of shoes, when a child next learns to *name* one particular shoe, her naming behaviour can extend to other shoes via the listener behaviour in her repertoire. Horne and Lowe point out that extension of the name relation can take place when new objects are physically similar to objects the child can already name, or when new objects serve a similar function. For the latter case, the authors give the example of a child putting her foot in a box, and then saying "shoe". The box evokes the listener behaviour that the child usually shows with regard to shoes, which in turn evokes extension of the name. Caregivers may correct extensions that do not coincide with the practices within the verbal community, thus shaping the child's naming repertoire in accordance with these practices.

In the initial stages of development, the child's listener behaviour repertoire will be far more extensive than her naming repertoire, but when she becomes more skilled in echoic behaviour this may have a strong effect on her naming repertoire, which may then expand rapidly. Also, whereas initially the child may need all the cues described in the examples above (the caregivers pointing, and prompting, etc.), and the component behaviours of the name relation may need to be learned separately, gradually the child needs less cues, and less repetition of hearing the object name in the presence of the same object, to learn new names. Children of 18 months old may learn to name new objects just by hearing someone else name them once or twice (e.g., Nelson & Bonvillian, 1973). At that point, Horne and Lowe note, "the name relation is established as a *higher order behavioural relation*" (1996, p. 203). Speaker and listener functions are then combined within the individual such that when the child is taught speaker behaviour, the corresponding listener behaviour is acquired as well, and when listener behaviour is taught, the corresponding speaker behaviour will be found to be in place too. Naming can then also become covert. Reinforcement provided by caregivers does not always remain as explicit as outlined above. It becomes more subtle but will still play an important role in the development of the child's naming repertoire (see Moerk, 1983, 1990, 2000).

There are common names for objects that have similar physical features (e.g., the names "dog", or "cat"), but also for objects that show little or no physical similarities (e.g., the names "toy", or "furniture"). In the latter case, one speaks of *arbitrary stimulus classes*. When the objects in a stimulus class have the same function (e.g., chairs), they form a *functional stimulus class* in which the objects are functionally equivalent (see Mead, 1934; and Goldiamond, 1962). That is, the stimuli in that class evoke the same response. So when a child who has learned appropriate listener behaviour with regard to chairs (i.e., to sit on them), hears that a novel object is a chair, she can then generalize her previously learned listener behaviour to this novel object, on the basis of the name. Alternatively, when she sees someone else show this particular listener behaviour with regard to a novel object, the child may then call it a chair. In this case the child generalizes the name on the basis of the listener behaviour, without being trained or reinforced on previous occasions.

At some point, a child learns that there are different levels of functional equivalence. For example, after having learned to name chairs, she later learns that the name "furniture" also applies, at what Horne and Lowe call "a more general level of functional equivalence" (1996, p. 205). And, as in the example with the chair above, when she hears this more general name "furniture" being used to refer to a new object (in her doll house), she can then show appropriate listener behaviour to the new object (e.g., when asked to put the furniture in a box) without being trained or reinforced to do so, on the basis of extension of the name. These behaviours that children show without being trained or reinforced to do so, are called *emergent* or *derived* relations. Horne and Lowe emphasize that these emergent behaviours, although not directly trained or reinforced, are still *learned* behaviours, established through naming. Moreover, naming itself is learned behaviour as well, and "its operations can be understood through behaviour analysis" (1996, p. 208). Emergent behaviours will also be discussed later, in relation to stimulus equivalence.

All of the above, then comes together in Horne and Lowe's definition of naming as "a higher order bidirectional behavioral relation that (a) combines conventional speaker and listener behavior within the individual, (b) does not require reinforcement of both speaker and listener behavior for each new name to be established, and (c) relates to classes of objects and events" (1996, p.207).

In addition, Horne and Lowe outline several ways to test whether a child, or any other subjects for that matter, can name. For example, by giving tact training with regard to novel stimuli, followed by testing for the corresponding listener behaviour. Alternatively, one can name the novel stimuli in the presence of the child, without reinforcing the child's behaviour, and then testing for tacting and for listener behaviour. And finally, one can train either listener or speaker behaviour with regard to a stimulus, and then extend the behaviour to new class members that are physically different, by saying the relevant common name in the presence of the new object(s). Following this, one can test the child by presenting the class members along with other stimuli, holding up one class member, and asking, "where are the others?" In short, this is giving the child a category-sorting test, after providing listener or speaker behaviour training separately for each class member.

Apart from the three component behaviours that together constitute naming, Horne and Lowe also discuss two other operants identified by Skinner (1957) – the intraverbal and the mand – and their relation to naming.

Intraverbal naming. Horne and Lowe (1996) write that around the age of 18 months children start to combine names. And they point out that the appearance of intraverbals in a child's verbal behaviour is an important step in her verbal development, and more specifically in her name relations. This is because of the role of these intraverbals in bringing about new or emergent behaviour. Initially, intraverbals can arise when a child hears her caregiver say /big teddy/ repeatedly, and she starts echoing this. The verbal stimulus /big/ can then come to evoke the child's verbal response "teddy", which is reinforced by the caregiver.

Horne and Lowe propose that there are intraverbals with, and intraverbals without conventional listener behaviour. When children learn nursery rhymes, for example, they show intraverbal behaviour without conventional listener behaviour. The authors emphasize that the latter form of verbal behaviour may be important for the further development of intraverbal naming (see also Skinner, 1957).

Apart from echoing name combinations uttered by caregivers or others, a child can produce name combinations herself too. For example, she may name two objects when seeing them together, or name an object and one of its properties (e.g., "doll", "pink"; or "helicopter", "flies"). Through self-echoic repetition these name combinations may become established as intraverbals, so that when the child produces one name (e.g., doll), hearing this name may evoke the other (pink). Self-echoic repetition may also make the intraverbal relations bidirectional (so that hearing /pink/ evokes saying "doll" as well as the other way around).

Horne and Lowe point out that once the child has learned intraverbal links of names, listener behaviour learned with regard to one of these names may transfer to the other name. How crucial this is, becomes clear in the example they give of the child who has learned to say "hot" when encountering a hot radiator, or any other hot object, and to show the appropriate listener behaviour of avoiding contact with it. Once the child has learned this, a caregiver can teach the child intraverbals such as "hot cooker" or "hot kettle", and both names that are paired with "hot" ("cooker" and "kettle") can come to evoke the same listener behaviour as the verbal stimulus "hot".

Naming and manding. When Horne and Lowe discuss manding and its relation to naming, they propose that this operant, just as the other operants, can occur with or without listener behaviour on the part of the speaker. It occurs in the absence of listener behaviour, for example, when a baby cries until she gets milk. And it occurs with listener behaviour when a child mands by saying "milk", after having learned to name milk.

Based on work by Anisfeld (1984), Halliday (1975), Terrace (1985), and others, Horne and Lowe note that "children generally learn to name objects and events before they learn to mand them. (…) names are first established and then are functionally extended to mand objects and events" (1996, p. 211). When names are extended to mand for something, the same speaker-listener functions and bidirectionality come into play as in normal naming. As for the establishment of manding, when a child has learned to name an object, she needs to be taught no more than a few mand relations, to enable her to extend newly learned names to mand (see Moerk, 1992; and Skinner, 1957).

Naming and verbal rules. With regard to naming and verbal rules, Horne and Lowe write that "names are the basic components of verbal rules" (1996, p. 213), and that therefore an analysis of the name relation is needed to get a better understanding of how rules affect behaviour.

Names can specify contingencies, but more importantly, they evoke listener behaviour to objects and events. Horne and Lowe describe how providing a (new) name for an object, can affect a child's listener behaviour towards the object. Their example is of a child who has a plastic bowl among her toys, and is told that it is a boat, or a hat. Having previously learned listener behaviour towards toy boats, and towards hats, and having learned to name boats and hats, the new name for the bowl can evoke listener behaviour appropriate to boats or hats (whichever name was supplied). Horne and Lowe note that the term *verbally controlled behaviour* (see Mead, 1934) covers this behaviour under the control of name relations better than Skinner's (1957) term *rule governed behaviour.*

Naming and stimulus equivalence. Horne and Lowe (1996) also describe how naming, and specifically common naming, can bring about stimulus equivalence. Particularly relevant in this context, are arbitrary stimulus classes (classes of physically different objects). Once a child has learned a common name for stimuli that look different, she may then, when shown one class member, be able to select the others without prior training.

Horne and Lowe give as example a child who learns the name “two” for the numeral 2, as well as for two dots (• •), and for the printed word TWO. As a result of learning this common name, an arbitrary stimulus class is formed, consisting of: 2, • •, and TWO. Now, when presented with either of these three stimuli, she can name the stimulus out loud, and hearing her own verbal stimulus, she can find the other class members in an array of stimuli in front of her, because the verbal stimulus she produced will evoke the previously learned listener behaviours of orienting towards them, and picking them up. Because the child has not received training to select, for instance, “2” when presented with “• •”, these are new behavioural relations, and Horne and Lowe describe them as emerging via listener behaviour.

As in this example, Horne and Lowe suggest that success on match-to-sample tests of stimulus equivalence might be due to naming of stimuli.

They point out that auditory-visual match-to-sample procedures, with an auditory sample stimulus and pictures or printed words as comparison stimuli, provide the right conditions for learning to name; one can echo the verbal stimulus, learn listener behaviour (picking the correct comparisons), and thus learn to name the comparison stimuli.

In visual-visual match-to-sample procedures using novel abstract stimuli, there may not be conventional names for these stimuli, but subjects may still find names for either the stimuli, or features of them. And if one finds common features to name for all members of a category, the stimuli would no longer be arbitrary.

Because naming is more easily established in the first procedure, subjects would be expected to learn baseline relations quicker, and do better on equivalence tests than in the second procedure. Studies by Green (1990), Lipkens, Hayes and Hayes (1993), and Sidman, Willson-Morris and Kirk (1986) support this expectation.

The naming that can help subjects to pass equivalence tests, can be common naming, in which one name applies to all class members, or it can be intraverbal naming, in which the subject gives the stimuli in a class different names and then learns intraverbals to remember which stimuli go together. For example, if a vertical bar should be matched with a green stimulus, the subject might learn the intraverbal "up-green" (see Lowe & Beasty, 1987). Horne and Lowe (1996) explain that through self-echoic repetition of this intraverbal, the names "up" and "green" can become bidirectionally related, and so can listener behaviour with regard to both stimuli (selecting one stimulus may then evoke selecting the other as well). Presented with one stimulus, the subject may name it, and hearing the name can evoke the intraverbally related name, which can in turn occasion the listener behaviour of selecting the comparison stimulus.

With regard to stimulus equivalence, Horne and Lowe (1996) put forward three predictions that arise from their account and are "open to experimental disconfirmation". The first prediction is that nonhuman organisms, that lack naming skills, will generally fail tests of stimulus equivalence tests. Horne and Lowe discuss relevant studies. So far, stimulus equivalence has not been demonstrated unequivocally in animals (see Schusterman & Kastak, 1993; but see also Horne & Lowe, 1996, p. 223-224). More relevant to the present thesis, with its focus on child development, are the remaining predictions stemming from the naming account. These are two predictions with regard to stimulus equivalence in humans, and Horne and Lowe describe related studies. The first prediction is that humans without naming skills will fail stimulus equivalence tests.

A study by Devany, Hayes and Nelson (1986) supports this prediction; normal 2-year-olds as well as mentally handicapped 2- to 4-year-olds with verbal skills passed, while mentally handicapped 2- to 4-year-olds without verbal skills failed equivalence tests in this study. Barnes, McCullogh, and Keenan (1990) also found that verbal ability and success on stimulus equivalence tests were related, in normal and hearing-impaired children. And Lowe and Beasty (1987) found that as age increased (from 2 to 5 years), which is of course linked to improvement of verbal skills, more children passed equivalence tests; all of the 4- to 5-year-olds, and half of the 3- to 4-year-olds passed, while only 1 of 7 of the 2- to 3-year-olds passed.

The second prediction that the naming account makes for stimulus equivalence in humans is that teaching subjects names for stimuli in match-to-sample procedures may have a strong effect on further test performance.

Dugdale and Lowe (1990) found support for this. In cases of test failure in normal 4- to 5-year-olds, they taught the children a common name for members of a stimulus class, which led to success on the equivalence test. Eikeseth and Smith (1992) conducted another study using common naming. Their subjects were 4 autistic children, aged between 3.5 and 5.5 years, who had language deficits. Emergent stimulus equivalence relations were found in 2 out of 4 children, after teaching them to name visual stimuli (Greek small letters and capital letters, and their printed names), whereas these relations had not been demonstrated by any of the children in a match-to-sample test conducted before the names were taught. Another child did not reach mastery level after being taught to name, but did perform above chance level. One child's performance did not exceed chance level. The results of the pre-experimental tests reported by Eikeseth and Smith showed that this boy's language skills were considerably less developed compared to the other children.

Lowe and Beasty (1987) taught intraverbal naming, that is, name links combining names for the sample stimuli and the comparison stimuli that were paired in baseline training (e.g., "up-green", and "up-triangle"), to the children who failed the stimulus equivalence tests. This led to success for 11 of the 12 children.

Dickins, Bentall, and Smith (1993) report a study with normal adults. All subjects received match-to-sample baseline training, but subsequently one group learned intraverbal naming for stimulus combinations that were in conflict with the classes defined within the experiment, while the control group learned intraverbals that were not related to stimuli in the task. Comparison of the performance of both groups on equivalence tests showed negative interference from the conflicting intraverbal name training.

Finally, Mandell and Sheen (1994) used letter combinations and symbol combinations as stimuli, and found that when pronounceable stimuli were used in the match-to-sample procedure subjects performed better and were quicker to form equivalence classes than subjects for whom unpronounceable stimuli were used.

Horne and Lowe (1996) conclude that naming the stimuli (in any of the ways described) can greatly enhance success in match-to-sample procedures and subsequent stimulus equivalence testing—that is, if the subjects' naming of stimuli is in line with the stimulus classes as defined within the experiment. This position has mistakenly been characterized as invoking "verbal mediating responses" for the establishment of equivalence, even recently (see Tomanari, Sidman, Rubio & Dube, 2006, p. 349), although Horne and Lowe stressed in their 1996 paper, that "naming should not be viewed as *mediating* the establishment of stimulus classes: Naming *is* stimulus-classifying behavior" (1996, p. 226-227).

Naming research

Since the publication of the naming paper in 1996, Horne, Lowe, and colleagues have conducted a range of studies to test the predictions of the naming account. The main focus of these studies was the suggestion, in the naming paper, that learning a common name (e.g., vek) for several arbitrary, or disparate, stimuli may establish a category relation between these stimuli. When presented with one category member and asked for "the others" in a category-sorting test, the child names the sample, and hearing the self-produced name should result in the listener behaviour of orienting towards all the stimuli she has learned to name as such (that is, all the veks). However, if the child only has a listener behaviour repertoire with regard to these stimuli, but cannot name them, the naming account predicts that she will not be able to categorise the stimuli. In the absence of naming, seeing the sample stimulus will not evoke listener behaviour towards the other category members. These predictions were tested in two studies.

Lowe, Horne, Harris and Randle (2002) taught 2- to 4-year-old children to tact three arbitrary stimuli as "zag", and three other stimuli as "vek", in pairwise training. Then, the children were given a category match-to-sample test. Six of the 12 children passed this test when presented with a sample stimulus and asked for "the others". The other 6 passed a second test, which prompted them to name the sample before selecting the other category members. Finally, a subset of children was tested to see whether the corresponding listener behaviour with respect to the arbitrary stimuli was in place, in the absence of training for this. All passed. Therefore, they had not just learned to tact, but they had learned common name relations.

Horne, Lowe and Randle (2004) taught 1- to 4-year-old children listener behaviour with respect to the same arbitrary stimuli as employed in Lowe et al. That is, in pairwise training they learned to select the correct stimulus upon hearing either /zog/, or /vek/. All children failed the subsequent category match-to-sample test. Following this, a tact test was administered, which 7 out of the 9 children failed. However, when they were then trained to tact the stimuli, 5 out of 7 children passed the category sorting test.

The results of both studies are in line with the predictions of the naming account that common naming is necessary for the establishment of category relations among arbitrary stimuli, and that common listener relations alone are not sufficient to bring about categorisation.

Another way of assessing whether a category relation is established, as described by Horne and Lowe (1996), is to teach the child novel behaviour to one category member, and then to test for transfer of function to other category members. Having previously learned to name the stimuli, the child will also name the stimuli during function training, whether required or not, and whether overtly or covertly. As function training proceeds, hearing the name of the stimuli will often precede production of the novel behaviour that is trained. This way, hearing the name comes to evoke the novel behaviour to that stimulus. In the category transfer of function test, when presented with the category members that were not involved in training, the child will name the stimulus, which will in turn evoke the appropriate behaviour. However, if the child only has a listener behaviour repertoire with regard to these stimuli, but cannot name them, the naming account predicts no transfer of function. In the absence of naming, seeing one of the other category members will not evoke the novel behaviour. These predictions were tested in two further studies.

Lowe, Horne and Hughes (2005) gave 1- to 4-year-old children pairwise tact training using six arbitrary stimuli (three "zogs", and three "veks"). In a subsequent test, all children demonstrated appropriate listener behaviour to the stimuli, so common name relations were found to be in place. Following this, they were trained to clap to one of the zog stimuli, and to wave to one vek. And when tested after training, all 9 children showed transfer of function to the other category members. That is, they would clap or wave appropriately to the stimuli that had not been involved in function training. Seven of the 9 children were then given an additional test for the corresponding listener behaviour. That is, the child was asked to select the correct stimulus in response to the experimenter either clapping or waving. All children passed. Next, four of these children were given the category match-to-sample test, and all passed. For 3 children further tact training was done to extend the stimulus sets to nine stimuli each. All showed transfer of function and correct category sorting with the extended sets.

Horne, Hughes and Lowe (2006) gave 1- to 4-year-old children listener behaviour training with respect to the same arbitrary stimuli as used by Lowe et al. (2005). In pairwise training the children learned to select the correct stimulus upon hearing either /zog/, or /vek/. When tested for the corresponding tacts, 10 out of 14 children passed, demonstrating that, for them, common name relations were in place. All 14 children were then given function training as in Lowe et al. (2005), and subsequently, the children who could name the stimuli either showed transfer of function, or listener behaviour with regard to the functions, or both. Three of these children were next given a category match-to-sample test, which they passed. However, the children who were able to show appropriate listener behaviour but could not name the stimuli, failed to show transfer of function, or correct category sorting behaviour in the respective tests that followed. Three of these children received additional tact training. As a result, they went on to pass the transfer of function test, and when two of them were given a category match-to-sample test, they also passed. Finally, three children were given additional listener behaviour training to extend the stimulus sets. All three passed a subsequent tact test, as well as transfer of function and category match-to-sample tests.

The results of both studies support the naming account, in that children are not able to categorise arbitrary stimuli in the absence of common naming; listener behaviour is not enough to establish categorisation.

Naming and levels of categorisation

The aim of this thesis is to explore the limits of the naming account, by studying how children learn to categorise at different levels, and how naming may bring this about. Subjects of study are 3- to 4.5-year-old children.

The arbitrary stimuli employed in these studies are eight newly designed alien animals. These aliens belong to four categories at a lower level, and to two categories at a higher level, in the same way as different types of reptiles, and different types of fish are all called "animals" at a higher level, and various types of fruits and vegetables are called "food" at a higher level. Nonsense names are used to designate the aliens at the different levels. The aliens are called *hib*, *feb*, *tor*, or *lup*, at a lower level, whereby each name designates two disparate stimuli. At a higher level they are called *zaag*, or *noom*. A brief outline of the studies will now be given.

In Study 1a, the children are taught to tact the alien animals at the lower level, and then tested for corresponding listener behaviour. Figure 1.2 shows the trained tact relations with solid arrows, and the tested listener relations with broken arrows.

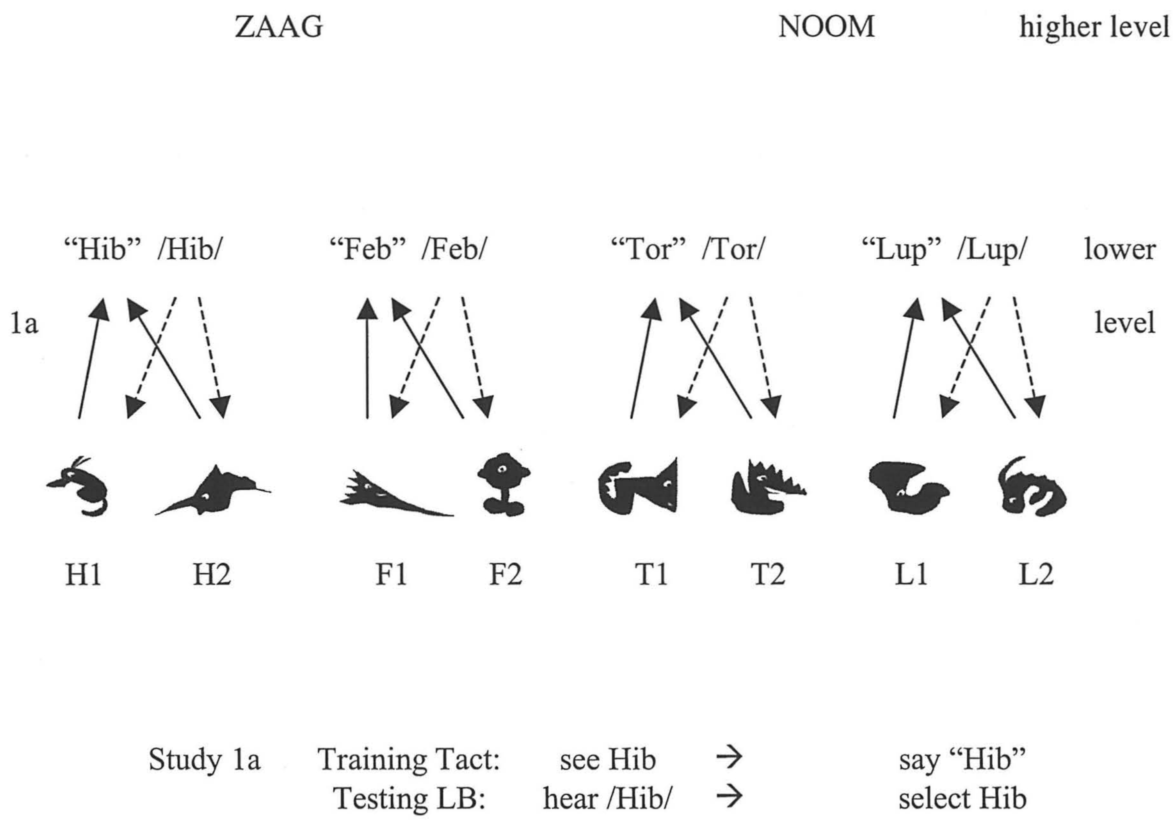
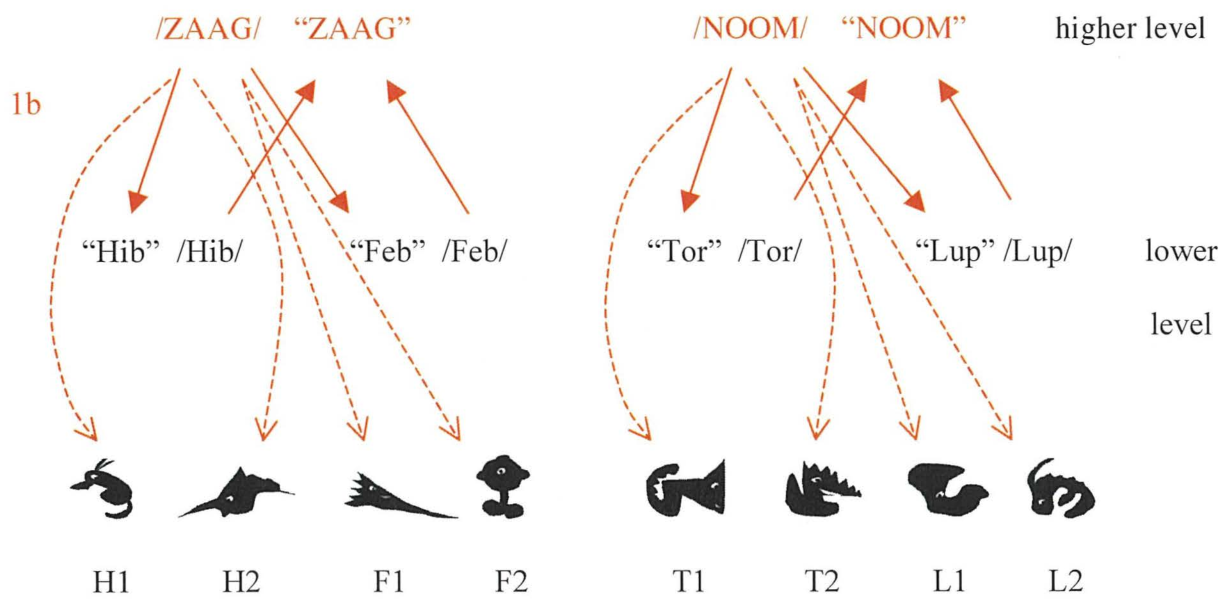


Figure 1.2 The trained and tested relations in Study 1a.

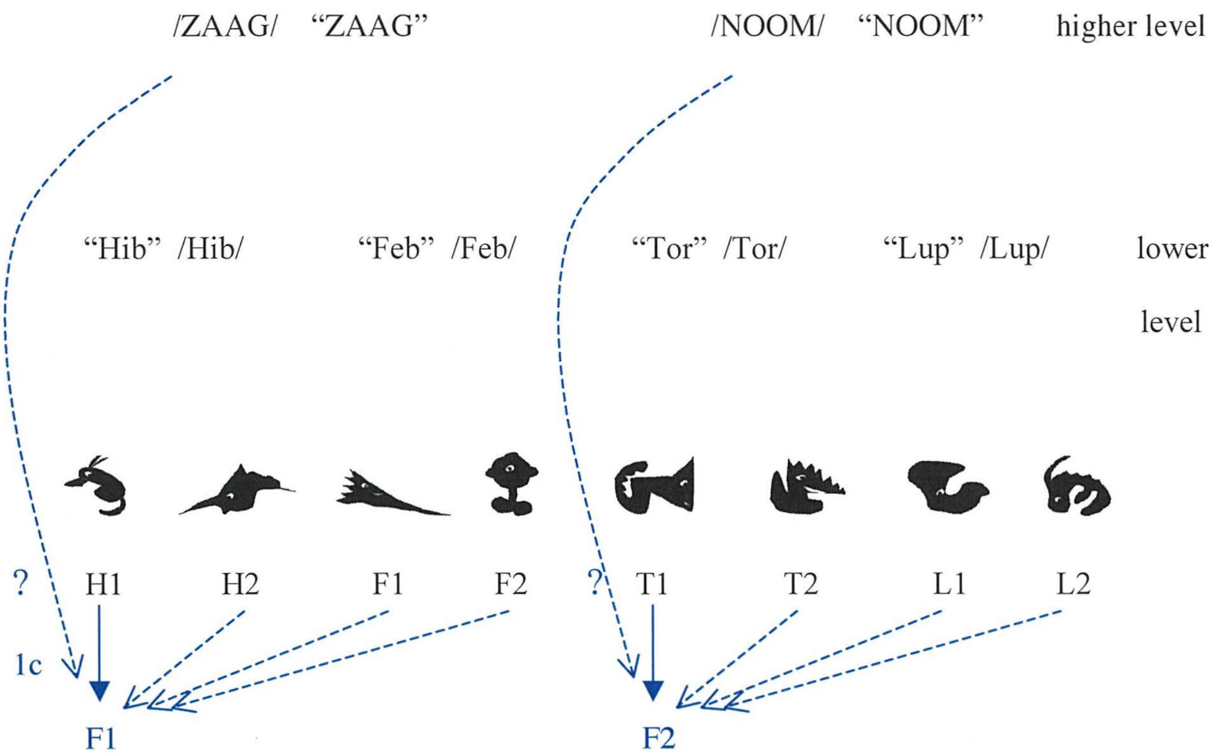
In Study 1b, the children learn intraverbals, relating the names of the aliens at the lower level with their potential names at the higher level, after which they are tested for listener behaviour at the higher name level. Figure 1.3 shows the trained and tested relations.



Study 1b	Training Intraverbal:	hear /Hib/	→	say "Zaag", and
		hear /Zaag/	→	say "Hib"
	Testing LB:	hear /Zaag/	→	select Hib

Figure 1.3 The trained and tested relations in Study 1b.

In Study 1c, a gesture (Function 1) is trained to one zaag, and another gesture (Function 2) to one noom. Next, the children are tested for transfer of these functions to the other aliens, via lower-level or higher-level names, or both (see Figure 1.4).

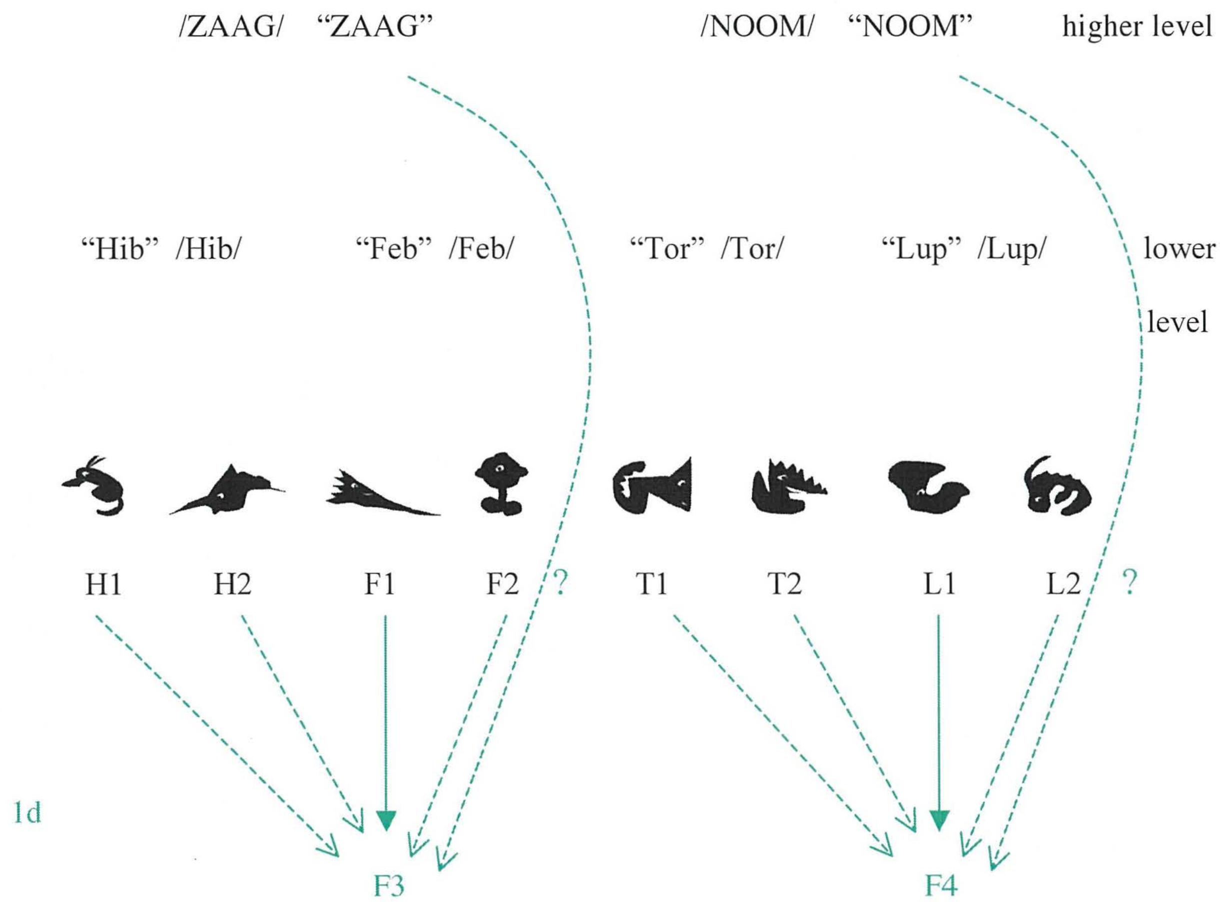


F1 / 2 = Function 1 / 2 (gestures).

Study 1c	Training F1:	see Hib 1	→	produce gesture (Function 1)
	Testing TOF:	see Hib 2	→	produce gesture (Function 1)? TOF via lower-level names?
		see Feb 1, 2	→	produce gesture (Function 1)? TOF via higher-level names?
	Training F2:	see Tor 1	→	produce gesture (Function 2)
	Testing TOF:	see Tor 2	→	produce gesture (Function 2)? TOF via lower-level names?
		see Lup 1, 2	→	produce gesture (Function 2)? TOF via higher-level names?

Figure 1.4 The trained and tested relations in Study 1c.

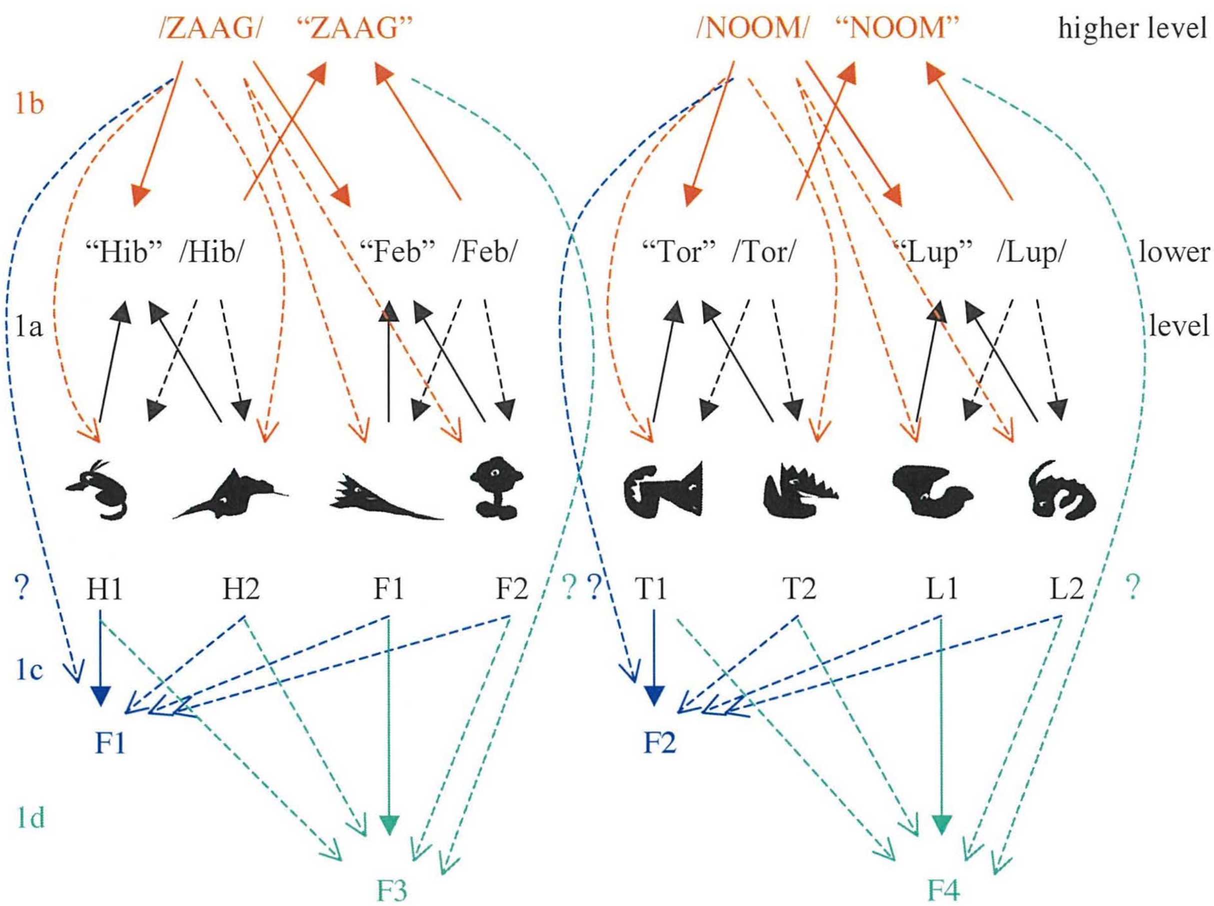
In Study 1d, one animal cry (Function 3) is trained to one zaag, and another (Function 4) to one noom. Then, a transfer of function test is conducted (see Figure 1.5). Figure 1.6 shows trained and tested relations of all four studies.



F3 / 4 = Function 3 / 4 (animal cries).

Study 1d	Training F3:	see Feb 1	→	produce animal cry (Function 3)
	Testing TOF:	see Feb 2	→	produce animal cry (Function 3)? TOF via lower-level names?
		see Hib 1, 2	→	produce animal cry (Function 3)? TOF via higher-level names?
	Training F4:	see Lup 1	→	produce animal cry (Function 4)
	Testing TOF:	see Lup 2	→	produce animal cry (Function 4)? TOF via lower-level names?
		see Tor 1, 2	→	produce animal cry (Function 4)? TOF via higher-level names?

Figure 1.5 The trained and tested relations in Study 1a.



Study 1a	Training Tact:	see Hib	→	say "Hib"
	Testing LB:	hear /Hib/	→	select Hib
Study 1b	Training Intraverbal:	hear /Hib/	→	say "Zaag"
	Testing LB:	hear /Zaag/	→	select Hib
Study 1c	Training F1:	see Hib 1	→	produce gesture (Function 1)
	Testing TOF:	see Hib 2	→	produce gesture (Function 1)? TOF via lower-level names?
		see Feb 1, 2	→	produce gesture (Function 1)? TOF via higher-level names?
Study 1d	Training F3:	see Feb 1	→	produce animal cry (Function 3)
	Testing TOF:	see Feb 2	→	produce animal cry (Function 3)? TOF via lower-level names?
		see Hib 1, 2	→	produce animal cry (Function 3)? TOF via higher-level names?

Figure 1.6 An overview of the trained and tested relations in Studies 1a-1d.

Study 1e, was a test only, investigating whether a more general verbal prompt than the prompts used in the tests of Studies 1c and 1d, would have stimulus control over both types of functional responses (gestures and animal cries) trained previously. The prompt for the gestures in Study 1c was "how does this one go?" while for the animal cries in Study 1d it was "what does this one say?". The more general prompt used in Study 1e was "what can this one do?".

Study 1f tested whether listener relations (i.e., see gesture – select animal) were in place for the gestures that the children were trained to produce in Study 1c (see animal – produce gesture).

Study 1g tested for listener relations (i.e., hear /animal cry/ – select animal) with respect to the animal cries the children were trained to produce in Study 1d (see animal – produce gesture).

Study 1h was a category match-to-sample test that investigated whether children were able to sort the alien animal stimuli into categories on the basis of the lower- and higher-level common names, without being trained to do so.

The second series of studies (2a-2h) are replications of the studies in the first series. The only difference was that in Study 2a, 2b, 2c, and 2h, multiple exemplar training with familiar stimuli (real life stimuli) preceded the studies with the alien animal stimuli. The procedure with the familiar stimuli was exactly as in the alien studies.

This experimental work on naming and levels of categorisation is described in further detail in Chapters 2 and 3.

Chapter 2 - Study 1

Study 1a – Learning to name

Study 1a focused on learning of naming, and on the relation between tact behaviour and listener behaviour. It tested the naming account, in that it investigated whether children can learn to tact eight different newly designed toy "alien" animals (i.e., see animal – say "name"), randomly allocated to four common tact categories with two members each. It also tested whether, without being directly trained to do so, the children demonstrated the corresponding listener behaviour on hearing the object name (i.e., hear /name/ – select animal).

Furthermore, the study investigated a novel way to train common name categories. The eight experimental stimuli were presented, four stimuli at a time, by placing them on a carousel. This allowed rapid presentation of stimuli, and the learning of simultaneous discriminations between all four stimuli presented (a necessary requirement for later categorisation tests). This method was compared in effectiveness (that is, in the number of trials to criterion) to procedures used in earlier Bangor research on categorisation, which entailed pair wise presentation of stimuli (see e.g., Lowe, Horne & Hughes, 2005).

Method

Participants

Participants were 13 children (6 females and 7 males) between the ages of 3 years and 1 month, and 4 years and 6 months at the start of the study (see Table 2.1).

They were recruited from Tir na n-Og, the Daycare Nursery and Centre for Child Development at the University of Wales Bangor. Initially 3 more children were recruited (2 females, 1 male). Due to major problems with staying on task, and consequent lack of progress, one female participant had to be dropped from the study after 20 sessions. Another girl (Belinda) was recruited to take her place. Another female participant, once in the experimental room, refused to say any alien name, even after modelling. After one session she was not willing to return. The male participant seemed distressed by the experimental situation in which he was asked questions, although in general he was very comfortable with the experimenter. After eight 12-trial blocks of training, in which he only responded to the initial prompt once, the sessions with him were discontinued. According to routine assessments conducted by the nursery nurses, the participants in Studies 1 and 2 showed no sign of developmental delay. For all studies, recruitment of the children was subject to parental consent, and ethics approval was gained from the School of Psychology Research Ethics Committee.

Table 2.1. Participants' sex, age at start of training, and age at first listener behaviour test.

Participant	Sex	Age at start (years/months*)	Age at testing (years/months*)
Davey**	M	3/1	3/5
Jon	M	3/4	3/5
Wendy	F	3/5	-
Lyn	F	3/5	3/8
Jim	M	3/6	3/7
Sebastian	M	3/6	3/7
Belinda	F	3/7	3/7
Ellie	F	3/7	-
Sara	F	3/7	3/9
Alun	M	3/8	3/10
Kyle	M	3/9	3/11
Cameron	M	4/6	4/9
Sasha	F	4/6	4/9

* age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).
** For reasons of confidentiality, these are not the children's real names.

Apparatus and Stimuli

The study was conducted in a purpose-built research room in Tir na n-Og Nursery (see Figure 2.1). The room was equipped with two wall-mounted video cameras and a microphone. One camera focused on the child, and the other on the experimenter, so as to produce split screen video recordings of child and experimenter in all sessions. A radio microphone enabled all sounds to be recorded. All recordings were controlled from the audio/visual suite (see Figure 2.2) located in a separate room in the nursery. The experimenter and the child sat opposite each other at a small table. The experimental stimuli were eight newly designed toy “alien” animals (see Figure 2.3). The aliens were made out of red Fimo (a thermosetting clay that was baked in the oven, after modelling of the required form) and were approximately 5 x 7.5 cm in size. All aliens had one eye. They were presented on a carousel, constructed from the lid of a biscuit tin (with a diameter of 20 cm) placed upside down, and with a small piece of cardboard attached underneath so that it could be rotated freely. Two identical carousels were used in the study, one for presenting the Set 1 stimuli and one for presenting Set 2. In no reinforcement sessions, a penguin hand puppet referred to as "Peter Penguin" was used.

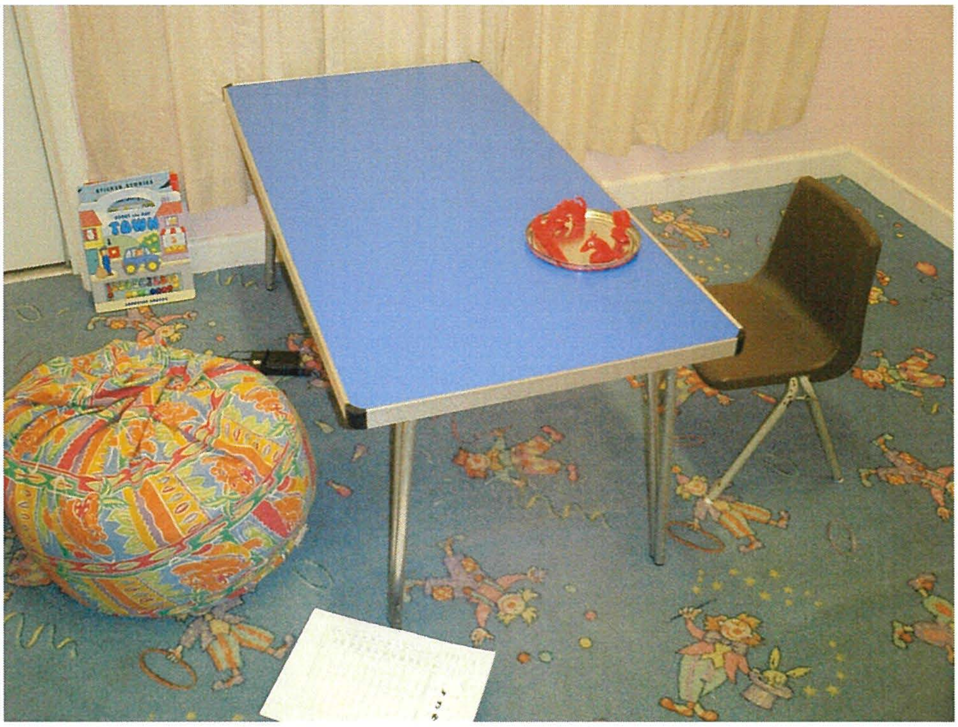


Figure 2.1 The experimental room

Figure 2.2 The audio/visual suite

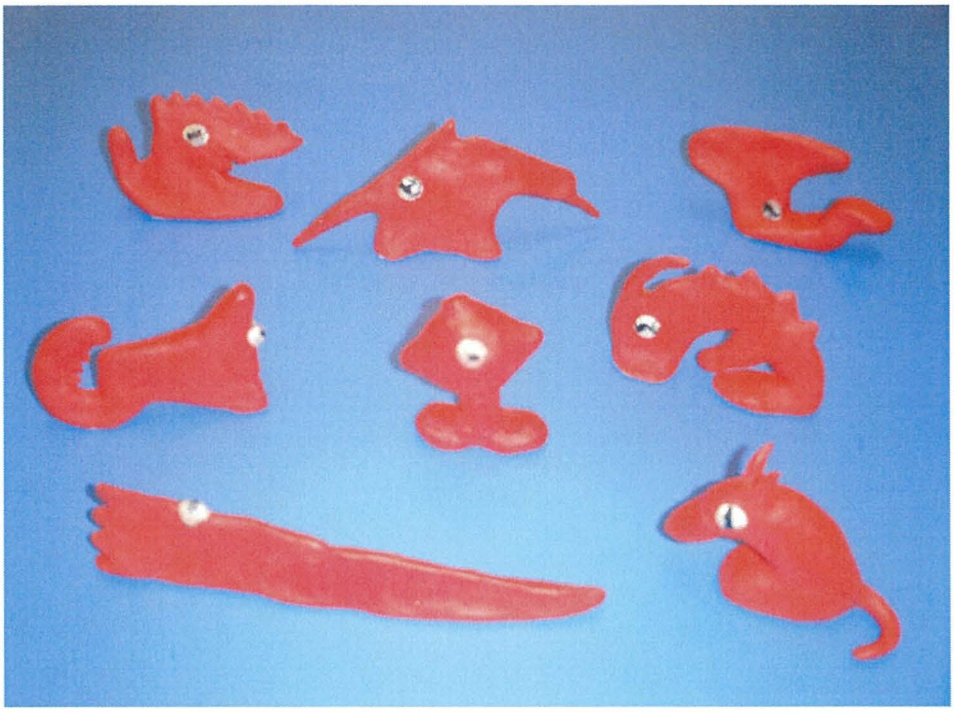


Figure 2.3 The experimental stimuli made out of Fimo

Stimulus sheets placed beside the table were used to schedule the presentation of the toy alien animals in each session. The child's responses were also recorded on these sheets during the sessions. To ensure that Experimenter 2 could not cue the child's responses during listener behaviour test sessions, the stimulus sheets that were employed specified the positions in which the aliens should be placed on each trial and the listener stimulus to be presented: Experimenter 2 was therefore unable to determine the name trained for each alien.

During test sessions, a wooden screen was placed on the table between the child and Experimenter 2 (see Figure 2.4). The screen was divided into two sections; the top section contained a clear perspex window that was covered, on the experimenter's side, by a net curtain through which the second experimenter could view the stimuli placed on the table and the child's responses. At the bottom section of the screen was an aperture that was partly covered by crepe paper. During testing, Experimenter 2 presented the stimuli to the child through the aperture.

Figure 2.4 The test screen

During training sessions, social praise was provided as reinforcement. In addition to this, when a child required a large number of trials to reach criterion, and appeared distracted from the task, a music book was used for reinforcement: the child could press a button for each, or every three, correct responses during the session. At the end of both training and testing sessions the experimenter read a story, which was selected by the child. The books were placed beside the table.

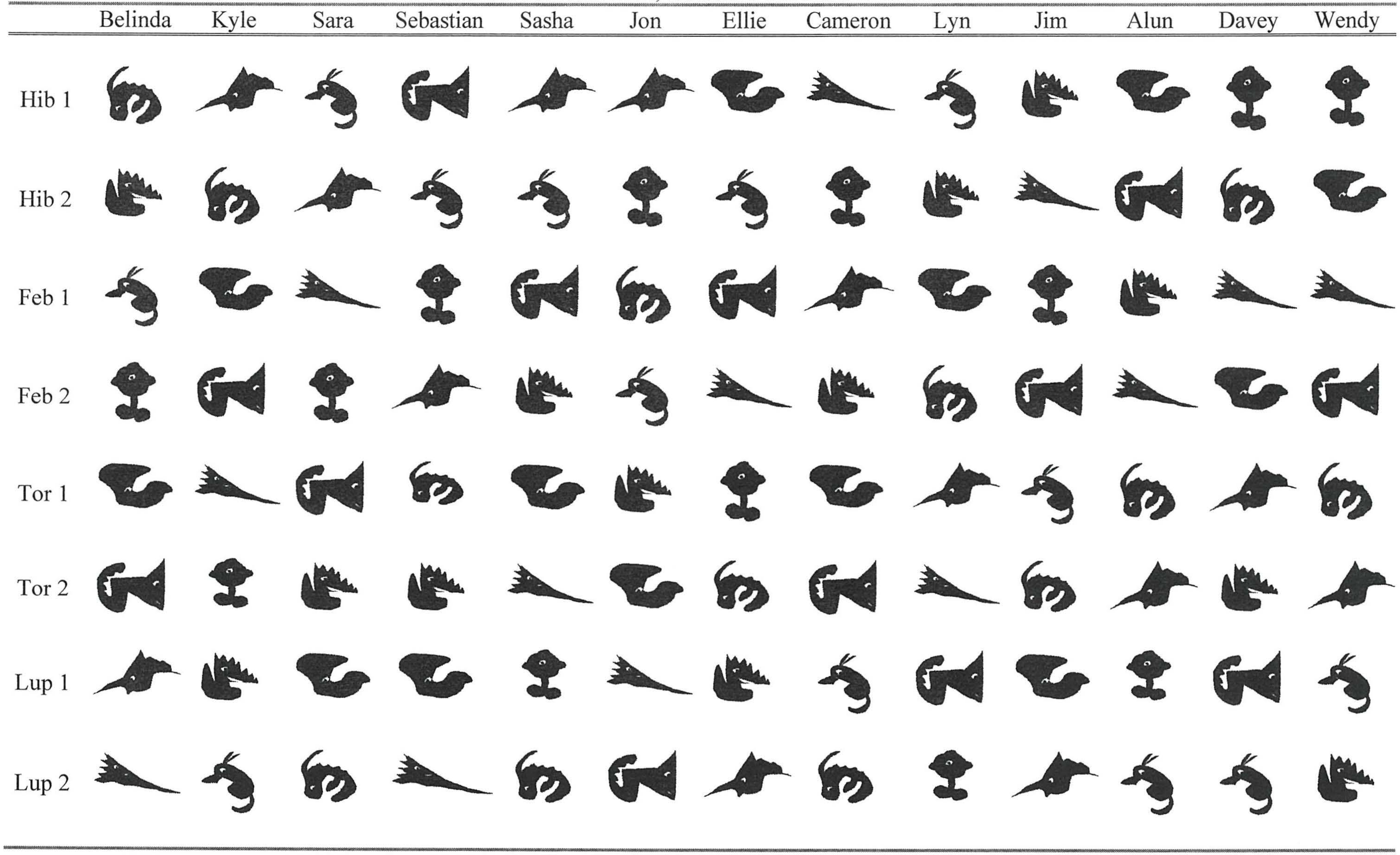
Procedure

A single case design with replication across participants was employed. Throughout the study there were no parents present. In all stages of the study there were one or two daily sessions with the child. Sessions varied in duration from 10 to 20 min.

Arbitrary stimulus tact training. First, the children were familiarised with Experimenter 1 during play in the common nursery rooms and the experimental room. Once the child was comfortable with the experimenter in both settings, the child was taken into the experimental room and shown the toy alien animals.

For each individual child, the eight stimuli were randomly assigned to four two-member categories, denoted by the nonsense names "hib", “feb”, "tor", and “lup”; these were further divided into two sets, each consisting of one hib, one feb, one tor and one lup. The specific animals and the names assigned to them for each child are shown in Table 2.2. An overview of the trained and tested relations in Study 1a is shown in Figure 1.2 (in Chapter 1).

Table 2.2 For each child, the names allocated to each alien.



The four Set 1 stimuli (Hib 1, Feb 1, Tor 1, Lup 1) were presented on a carousel, in relative positions that were counterbalanced over trials. Before starting a trial, the experimenter made sure that the child was attending. At the start of each trial the experimenter, or the child, would spin the carousel around. The experimenter stopped the carousel when the target stimulus was directly in front of the child. On the first trial for each alien, the experimenter pointed at the target stimulus and introduced it by saying, “Look at this, it’s a hib/feb/tor/lup. Can you say hib/feb/tor/lup?” On every subsequent trial with the same alien, the experimenter pointed at the targeted stimulus and said, “What’s this?” Following a correct response the experimenter delivered social praise (saying e.g., “Yes, well done!”). And occasionally, in addition to this, a music book was presented on which the child could press a button. If the child did not produce the correct response, the experimenter provided corrective feedback: “It’s a hib/feb/tor/lup. Can you say hib/feb/tor/lup?” One stimulus was targeted on each trial. The aliens were each targeted three times in a pre-specified quasi-random order in blocks of 12 trials in which the same trial type did not appear twice in succession. At the start of each new block of 12 trials, the toy animals were placed in a different position on the carousel. The learning criterion for this stage was 10 out of 12 correct over two consecutive 12-trial blocks. Following this procedure, the child was trained to criterion in the same manner, with Set 2 (Hib 2, Feb 2, Tor 2, Lup 2). During training for Set 2, maintenance training trials were conducted with Set 1. Maintenance training continued until the child produced 100% correct responses over four trials.

After the child had learned to tact the aliens in both sets, mixed arrangement training took place, to ensure that correct tacting behaviour was maintained also when stimuli were placed in a context other than the one in which the initial training had occurred.

Mixed sets. Either the hib (H), the feb (F), the tor (T) or the lup (L) in Set 1 was replaced by the alien of the same category in Set 2, and vice versa. If in mixed arrangement 1, for example, the hibs were exchanged, the stimuli arrangements were: H2, F1, T1, L1 and H1, F2, T2, L2. After reaching criterion on this arrangement, the particular toy animals were placed back in their original set and two other animals were switched around (e.g., the febs were exchanged). The interchanging took place in a random order for individual children, and continued until the animals in all four two-member sets had been switched around, such that the child demonstrated correct tacting of each alien in the presence of all aliens from the alternative categories.

Reduction in reinforcement rate. In the very last training stage, with randomly generated mixed sets consisting of two stimuli from Set 1 and two from Set 2 (like Sets 1 and 2, the mixed sets consisted of one animal from each category — e.g., H2, F1, T1, L2), reinforcement was reduced to 0%. This was done to make sure that the child's tacting behaviour remained in place in the absence of reinforcement, in preparation for the listener behaviour test in which there would be no contingent reinforcement delivered for responding correctly. In these no reinforcement sessions the experimenter introduced a penguin hand puppet (Peter Penguin), and asked the child to teach Peter the names of the aliens. The experimenter said she would keep very quiet, and wouldn't say, "Yes, well done!" all the time, because that would distract Peter. If the performance of any child did not meet the zero reinforcement criterion, then reinforcement was reintroduced until responding was once again at criterion. The child was then tested once again in extinction, and so on, until the child's performance met the zero reinforcement criterion.

Listener behaviour test. Following tact training, the child was tested for listener behaviour (hear name – select animal) discriminations. Prior to testing, Experimenter 1 conducted eight review tact trials, four with Set 1, and four with Set 2. Each stimulus was targeted once, to check whether the trained tact behaviour was still in place. After these trials, Experimenter 1 took a seat behind the child, while Experimenter 2 sat opposite the child. A screen was placed between Experimenter 2 and the child (see Apparatus and Stimuli), so that the experimenter's line of gaze could not cue the child's responses. Four of the alien animals (a random selection of two stimuli from Set 1 and two from Set 2, one animal from each category) were put on the carousel in pre-specified positions. Experimenter 2 tested the child's listener behaviour by spinning the carousel, and when it came to rest, asking the child: "Which one is the hib (or feb/tor/lup)?" If the child did not respond within 4 s, the verbal prompt was repeated, at most twice. Each stimulus in Mixed Set 1 was targeted three times in a 12-trial block, in a pre-specified quasi-random order in which the same trial type did not occur twice in succession. Before starting a second 12-trial block, the positions of the stimuli on the carousel were changed. After two blocks of trials with Mixed Set 1, the procedure was repeated with Mixed Set 2. The overall listener behaviour test criterion was 14 out of 24 correct for each mixed stimulus set (24 trials); the binomial probability of 14 or more correct out of 24 by chance is .0005. An additional criterion for each stimulus was set at 4 correct out of 6; the binomial probability of 4 or more correct responses by chance is .037. No reinforcement was given during the test.

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 99%. Similarly, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Tact training: Set 1, Set 2, mixed sets, and reinforcement reduction. Figure 2.5 shows the total number of trials for each individual child. Eleven of the 13 participants completed tact training. Of the 2 children who did not complete training, Wendy had a total of 900 trials but did not reach criterion for Set 2, while Ellie had 807 trials but did not reach criterion on the mixed sets at the no reinforcement stage. Both girls were given additional sessions[1], but left the nursery before tact training could be completed. The children who completed training needed on average 598 trials (range 372—984). To break this down for the separate stages, for Set 1 the average was 276 (range 168—396), for Set 2 it was 132 (range 36—432), for the mixed sets with reinforcement it was 140 (range 96—228), while for the mixed sets

---

[1] Ellie was given additional sessions in which only the aliens she kept making errors on were targeted (Hib 2 and Feb 1), plus a few trials in which the experimenter and Ellie would say the name of the alien together in the presence of the animal.

In the additional sessions with Wendy, only Feb 1 and Lup 1, and Feb 2 and Lup 2 were targeted. When this did not result in improved performance, these animals were again presented in pairs of one lup and one feb, but with a golden star underneath Lup 1 and Lup 2. The child was asked under which animal the star could be found and she succeeded on this task. Then she was told the "secret" that the one with the star underneath was the lup. In the trials that followed, she was first asked "Where's the star?" then, when she picked up an alien, "what is it?" and finally, the experimenter pointed at the other animal and asked "and what's this?" When a normal 12-trial block was conducted with the full Set 1 she correctly tacted the lup (three times), but not the febs. Pair wise training was then resumed with the lups and febs. The lups still had the star underneath, and the febs had a golden triangle underneath. This procedure did not result in correct tacting. Wendy once again showed poor task compliance.

no reinforcement it was 50 (range 24—72). When it comes to the total number of trials to complete training, the 6 females needed on average 615 trials (range 528—732), while the 7 males needed 588 (range 372—984). In Figure 2.5 the numbers of trials are presented as a function of age, showing that the older children did not reach criterion in less trials than the younger ones.

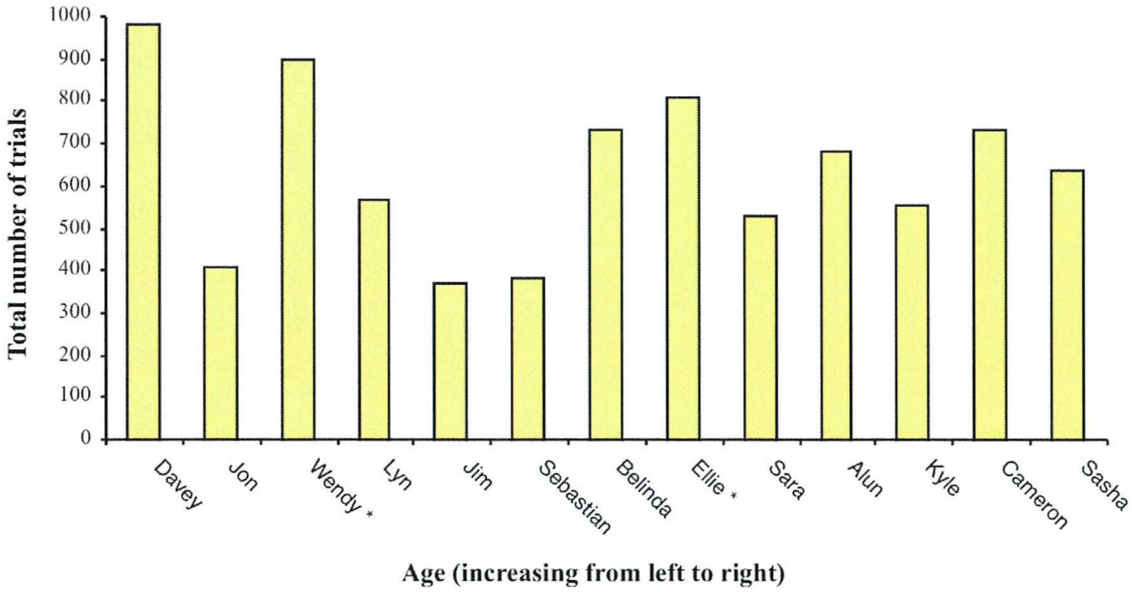


Figure 2.5 The total number of trials needed to reach criterion over all stages until the (first) listener behaviour test, for all children with the youngest child on the left (Davey, 3/1) and the oldest child on the right (Sasha, 4/6). For all ages: see Table 2.1. (*Child did not complete the study.)

Figure 2.6 shows the number of trials for female and male participants separately. Though on average the males needed fewer trials than the females, this difference was not consistently reflected in the individual data.

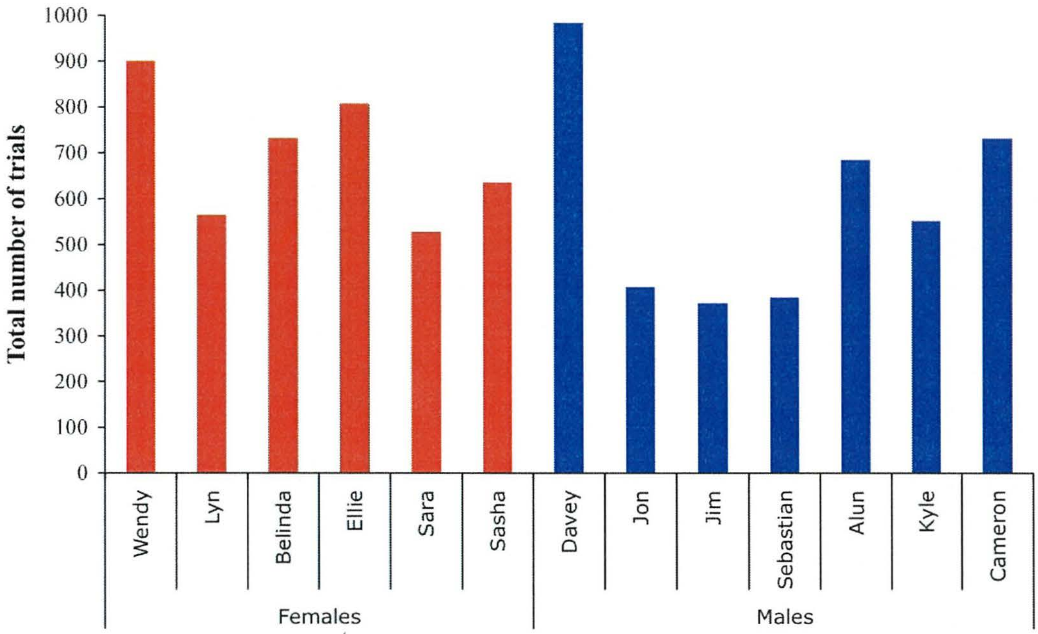


Figure 2.6 The total number of trials for females (in red) and males (in blue).

Figure 2.7 shows the breakdown of the number of trials to criterion for each stage within the study and for all individual children. It can be seen that out of 13 participants 11 needed considerably less trials for Set 2 than Set 1. When presented with a second set of aliens, the child was already familiar with the procedure, and the names (both sets consisted of animals named hib, feb, tor, and lup). For Set 1, none of the participants reached criterion in the minimum number of 36 trials. For both Set 2 and the mixed sets, Lyn only required the minimum of 36 and 96 trials, respectively, while all other children needed more trials. At the mixed sets with no reinforcement stage, 8 children required no more than the minimum of 48 trials.

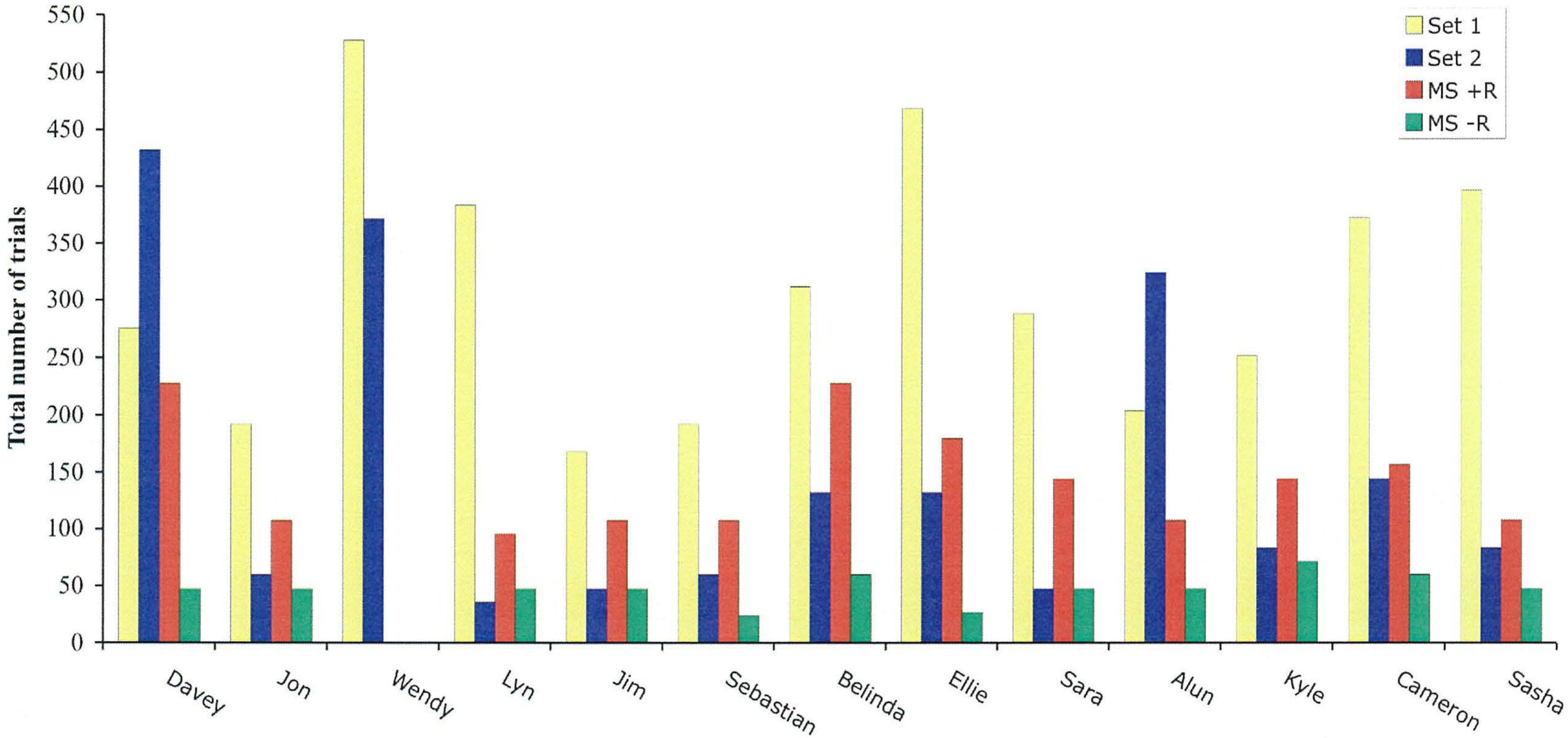


Figure 2.7 For each child, the total number of trials to criterion for Set 1, Set 2, mixed sets with reinforcement (MS+R), and mixed sets without reinforcement (MS-R).

Listener behaviour test. Figure 2.8 shows the percentage of correct responses per child for each of the two mixed sets. The pass level for the 24-trial block for each set was 58%. As Figure 2.8 shows, all 11 children passed the listener test. Four children made no errors in the total of 48 trials (there were 24 trials for each of the two mixed sets), 4 children made one error, and 1 child made two errors. However, the 2 remaining children (Davey and Belinda) showed more variable performances, particularly on Mixed Set 1. Davey made 10 errors on Mixed Set 1, and two on Mixed Set 2, and Belinda made four errors on Mixed Set 1, and three on Mixed Set 2. Davey met the individual stimulus listener criterion for two of the four stimuli in Set 1 (but not for Hib 1 and Tor 1) and all stimuli in Set 2. After the first test, he was given more tact training, including sessions that only targeted the two animals for which he showed consistent errors (Hib 1 and Tor 1). Three more listener re-tests were then conducted. Although he reached above chance performance for Hib 1 on the second and fourth re-tests, his performance on Tor 1 trials remained below chance.

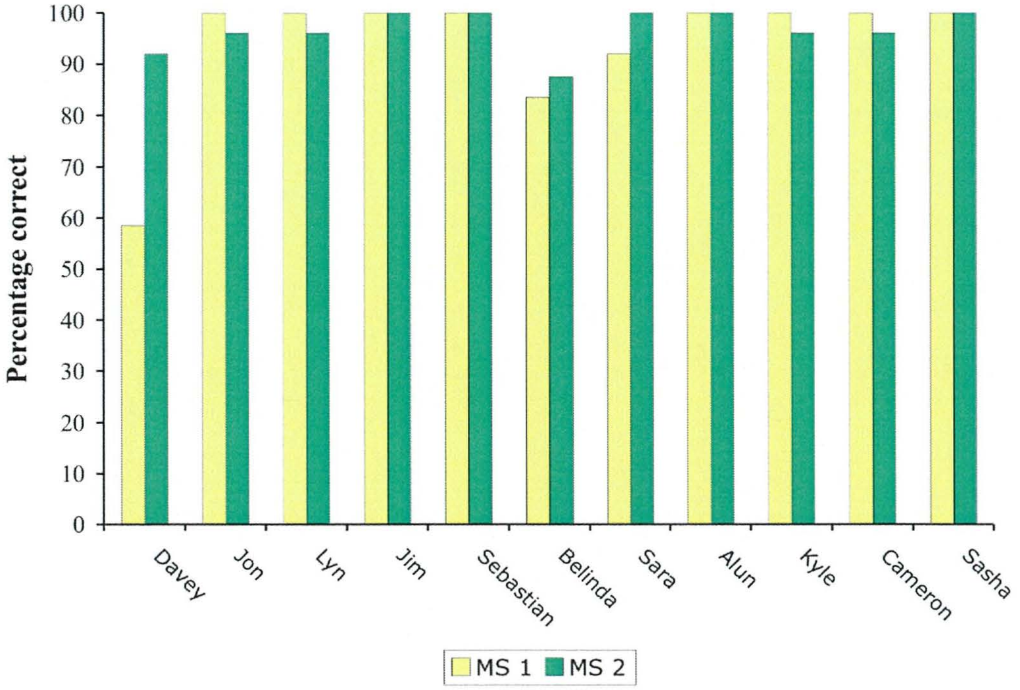


Figure 2.8 For each child, the percentage of correct responses in the listener behaviour test for Mixed Set 1 (in yellow), and Mixed Set 2 (in green). Pass level for each set was 58%.

Table 2.3 lists the vocalisations produced by the children during the listener behaviour test at the lower name level. Appendix C contains a collection of comments that the children made during training sessions of all studies.

Table 2.3 The children's vocalisations during the listener behaviour test at the lower name level.

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Davey	Mixed Set 2	Which one's the feb?	(points at tor) Tor. (Then points at feb.)
Alun	Mixed Set 2	Which one's the feb?	(points correctly) There! He's looking at you. Naughty feb. He doesn't want to play.
Cameron	Mixed Set 2	Which one's the lup?	Lump (This is what he consistently called the lup.)
Sasha	Mixed Set 1	Which one's the feb?	Feb, feb, feb. (Then points correctly.)

Discussion

The first aim of Study 1a was to see whether children could learn to tact (i.e., see animal – say "name") the eight newly designed "alien" animals. The data show that 12 out of 13 children reached criterion for tacting the aliens in both Set 1 and 2, while 11 children reached criterion for all stages in the tact training procedure. Therefore, the alien animals have been shown to be suitable experimental stimuli, which could be useful in further studies.

The second aim of this study was to investigate whether children who had learned to tact the alien animals, would demonstrate the corresponding listener behaviour on hearing the object name (i.e., hear /name/ – select animal), without being directly trained to do so. As mentioned above, 11 of the 13 participants reached criterion for all stages of tact training and were tested for listener behaviour; all 11 passed the listener behaviour test the first time, thereby providing support for the naming account (Horne & Lowe, 1996), which suggests that when a child is trained to tact an object, in the course of this the child is likely to also learn the corresponding listener behaviour, without it being separately trained. This would apply to normally developing children of 2- to 4-years old (see Lowe, Horne, Harris & Randle, 2002). The participants in Study 1a are 3- to 5-years old. Davey was the only child to fail the individual stimulus test for listener behaviour for one of his stimuli (Hib 1) and he scored just above chance for another stimulus (Tor 1). However, the errors he made in the listener behaviour test involved names and animals he also had great difficulty with in his tact training. In total he needed 984 tact trials because he continually made the same error. It was with the same stimuli that he made errors on the listener behaviour test. Although Davey's tact performance eventually met the criterion for tact training before each of three re-tests for the corresponding listener behaviour, the above would suggest that both his speaker and his listener behaviour for these particular stimuli were unstable. A further interesting observation is that in some trials, some of the children echoed the listener stimulus out loud before producing the listener response.

The third aim of Study 1a was to evaluate this novel way of presenting four stimuli at a time (by placing them on a carousel), thereby allowing the learning of simultaneous discriminations between all four stimuli. This new method can now be compared in effectiveness (that is, in the number of trials to criterion) to methods used in earlier Bangor research on categorisation, which involved pair wise presentation of stimuli (see e.g., Lowe, Horne & Hughes, 2005). When using eight animals in a study with pair wise presentation of stimuli, 24 combinations would be needed. In order for the child to learn to distinguish H1-F1-T1-L1 and H2-F2-T2-L2, which are the two initial sets of animals on the carousel in the present study, with pair wise presentation it would be necessary to train the following 12 combinations:

H1-F1	H2-F2
H1-T1	H2-T2
H1-L1	H2-L2
F1-T1	F2-T2
F1-L1	F2-L2
T1-L1	T2-L2

And after training these combinations, which would equal training of the two sets of animals on the carousel, it would then – in line with the stage of mixed sets in the present study - be necessary to train the following 12 combinations in pair wise training:

H1-F2	T1-H2
H1-T2	T1-F2
H1-L2	T1-L2
F1-H2	L1-H2
F1-T2	L1-F2
F1-L2	L1-T2

This adds up to 24 combinations. In the case of ideal learning, meaning errorless learning (which would be highly unlikely), for each of the listed combinations there would be one 8-trial block for introduction of the animals, followed by two blocks with 100% reinforcement (criterion 7/8 correct over 2 blocks) and one block without reinforcement (same criterion). Each stimulus would be targeted four times per block. So in total there would be 32 trials (4 blocks x 8 trials) per combination of 2 animals. Because there were 24 combinations of aliens, in the case of ideal learning 768 trials would be necessary in total for 8 aliens with pair wise presentation. In Study 1a, the child with the slowest learning process required 984 trials (82 blocks x 12 trials). Recall that there were 11 children in Study 1a who completed tact training. Ten of them required less than 768 trials; 6 required less than 600 trials; 2 required around 400; and the child with the fastest learning required only 372 trials. In short, the new method of presenting four stimuli at a time generally required considerably less trials, and is therefore far more effective than pair wise presentation.

Study 1b – Learning higher-level names

Objects and entities in the world have names at different levels of abstraction. Dogs and cats, for example, are also called "animals", at a higher level. Study 1b investigated how children learn higher-level names for objects and entities. Do they have to be explicitly told the higher-level names in the presence of the objects or entities, or is it perhaps enough for them to just hear the lower-level and higher-level names together?

To investigate this, a word game was introduced in Study 1b, for the children who had already learned the name relations with regard to the aliens in Study 1a. In this word game, the four alien names (hib, feb, tor, lup) were linked to two different higher-level names (zaag, noom). Just as dogs and cats are all "animals" at a higher name level, the names zaag and noom were used to designate the aliens at a higher level. For example, the hib and the feb could be called zaag at a higher level, and the tor and the lup could be called noom at a higher level. The children were trained the intraverbal relation between the lower-level and higher-level names. No alien animal stimuli were present during the word game. Once the children had learned the intraverbal name relations, they were tested for listener behaviour with regard to the higher-level names (i.e., hear /higher-level name/ – select animal).

Method

Participants

Of the participants in Study 1a who passed the test for listener behaviour, 8 children (3 females and 5 males) were available for part 1 of Study 1b. They were between the ages of 3 years and 7 months, and 4 years and 10 months at the start of the study (see Table 2.4). For part 2 of this study, Sebastian, Sara and Cameron were not available, because they had left the nursery.

Table 2.4 Participants' sex, age at start of training, and age at listener behaviour test, for Leg 1 and 2.

Participant	Sex	Age at start training Leg 1*	Age at testing Leg 1	Age at start training Leg 2	Age at testing Leg 2
Jon	M	3/7	3/7	3/11	3/11
Lyn	F	3/8	3/9	3/10	4/4
Jim	M	3/9	3/9	3/10	3/10
Alun	M	3/11	3/11	4/0	4/1
Sebastian	M	3/11	4/0		
Sara	F	4/0	4/0		
Cameron	M	4/9	4/9		
Sasha	F	4/10	4/10	4/11	5/2

* Ages in years/months; age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).

Apparatus and Stimuli

The apparatus and stimuli employed in Study 1b were as in Study 1a, except that there were no alien animal stimuli on the table during word game trials. Stimulus sheets listed the verbal stimuli to be provided by the experimenter, and were used to record the child's behaviour during the session. In the listener behaviour test session, the alien animal stimuli (see Study 1a) were used. In the test session, Experimenter 2 used stimulus sheets that specified the position of the aliens on the table and the order of targeting without linking the higher-level names and the shapes, so that she remained blind to the trained intraverbal relations and could not cue the child's responses during the test trials.

Procedure

This was a single case design with replication across participants. During the sessions no parents were present. There were one or two sessions per day with the child, taking about 10 to 20 min each.

Arbitrary stimulus tact training. For each participant, Study 1b was started as soon as possible after completion of Study 1a. Where necessary, the tact relations established in Study 1a were retrained to criterion.

Echoic training. In a word game the names of the aliens at the lower level (hib, feb, tor, and lup) were linked to the name of the animal at the higher level (zaag or noom). Counterbalancing took place with regard to which of the higher-level names was linked to which of the lower-level names for the individual children. So a hib could be a kind of zaag for one child, while it was a kind of noom for another child. Figure 2.9 shows the hierarchical structure with the intraverbal name links for 2 participants (Sara, at the top, and Sebastian, at the bottom). For Sara, the intraverbals first trained in the word game were hib-zaag and tor-noom, for Sebastian these were feb-zaag and lup-noom.

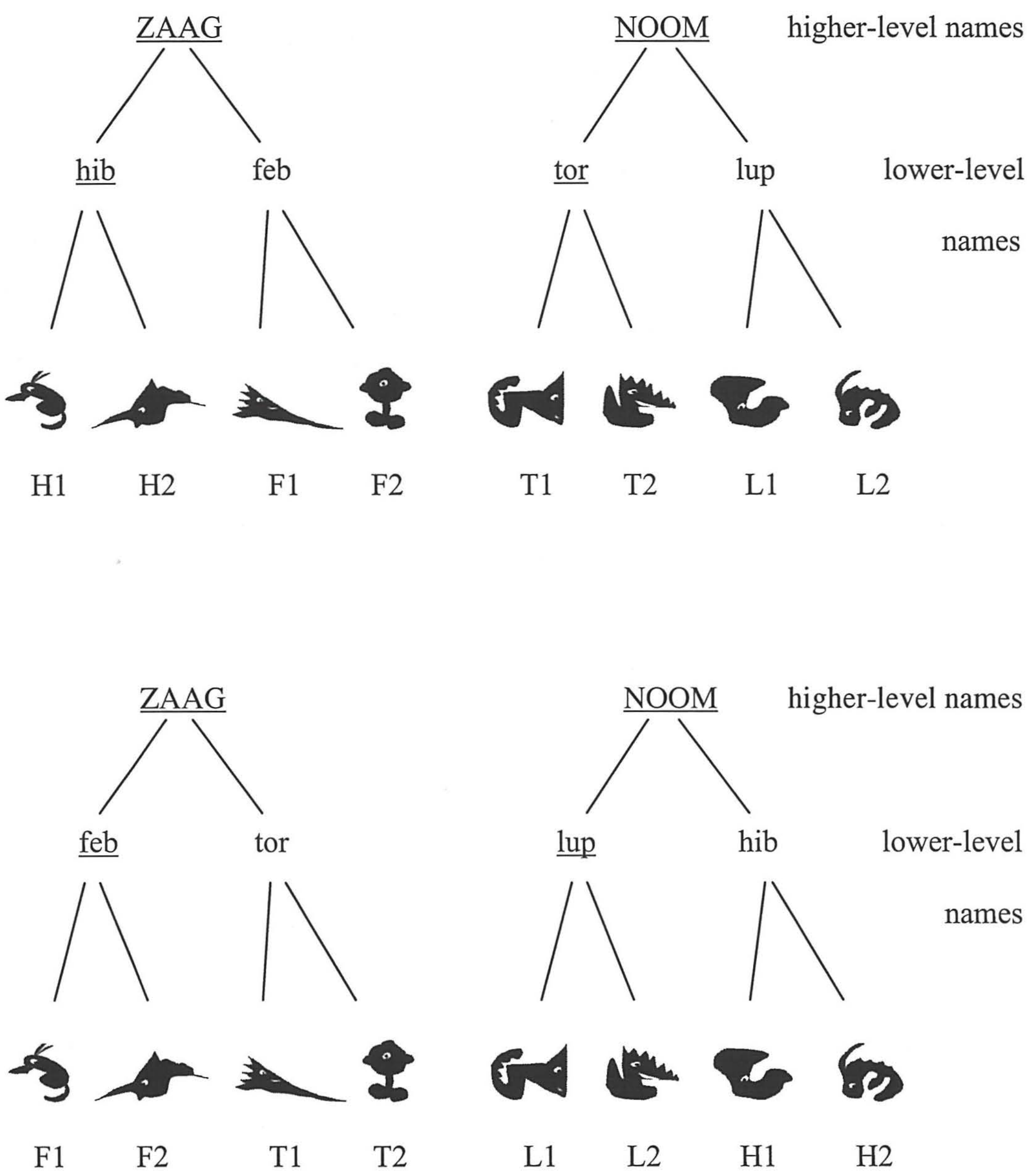


Figure 2.9 The hierarchical structure for 2 participants (Sara, at the top, and Sebastian, at the bottom), with the names of the aliens at the higher level and the lower level. Underlined are the links that were trained first in the word game. For Sara, these links were hib-zaag and tor-noom, for Sebastian these were feb-zaag and lup-noom.

The first part of the word game that linked the names of the aliens at the two levels, was an echoic game, called *You say what I say*. In this game the experimenter produced a verbal stimulus and the child produced an echoic response. The verbal stimulus became more complex over trials. That is, if the verbal stimulus in the first trial of a block of eight trials was, for example, “hib”, then it would be “hib-zaag” in the second trial. And in the remaining six trials, it would be “hib-zaag-hib-zaag”. In a second block of eight trials the verbal stimulus in all trials would also be “hib-zaag-hib-zaag”. The child’s responses on each trial were recorded on the stimulus sheets.

The child’s correct echoic responses were followed by social praise (e.g. “Yes!” or “Well done”). If the child did not produce the correct echoic response, the experimenter provided corrective feedback by producing the verbal stimulus again. The learning criterion in this part of the game was seven out of eight correct responses within a block over two consecutive blocks.

Intraverbal training. The second part of the verbal training was an intraverbal game, in which the child responded with a “chain” of lower-level and higher-level names upon hearing just the lower-level name. The child was instructed that whenever the experimenter said, for example, “hib”, the child should say “hib-zaag-hib-zaag”. In this part of the game, the experimenter provided the same verbal stimulus over a minimum of two blocks of eight trials each. Praise or correction followed the child’s response as described above. The learning criterion was, again, seven out of eight correct within a block over two consecutive blocks. The trained and tested relations in Study 1b are shown in Figure 1.3 (in Chapter 1).

Reduction in reinforcement rate. Once criterion was reached for the second part of the word game, a further two blocks of eight trials were conducted. The verbal stimuli in these trials were identical to the stimuli in the intraverbal game (the stimulus being, for example, “hib” and the expected response “hib-zaag-hib-zaag”). Now, however, the trials were presented under zero reinforcement. This was done to prepare the child for the zero reinforcement situation of the test sessions, and thereby to ensure that responding would continue also under those extinction conditions. In these no reinforcement sessions the experimenter reintroduced the penguin hand puppet (Peter Penguin), and asked the child to teach Peter the word game. The experimenter said she would keep very quiet, and wouldn't say, "Yes, well done!" all the time, because that would distract Peter. At this stage, as before, the learning criterion was seven out of eight correct within a block over two consecutive blocks. If the performance of any child did not meet the zero reinforcement criterion, then reinforcement was reintroduced until responding was once again at criterion. The child was then tested once again in extinction, and so on, until the child's performance met the zero reinforcement criterion.

All three stages in the word game that are described above were played with respect to both higher-level categories. Counterbalancing took place for starting the intraverbal training with either the zaag or the noom higher-level category names. For example, Sara started with the zaag link, while Sebastian started with the noom link. During the word game training period, maintenance training was conducted with Set 1 and 2 of the aliens. There were no alien animal stimuli on the table during the word game.

Listener behaviour test for higher-level names. The listener behaviour test determined whether echoic and intraverbal training that establishes bi-directional links between tacts/names at the lower level with potential tacts/names at the higher level resulted in the child being able to make listener behaviour discriminations (hear name – select animal) with regard to the higher-level names. In the test sessions, Experimenter 1 presented review trials for the lower-level tacts in two 4-trial blocks, one for Set 1 and one for Set 2. Following that, two 4-trial blocks were conducted to review the intraverbal links, one block for each link. During the review trials, occasional errors were corrected. After the review trials and before the test, Experimenter 1 gave the child the instruction, "Now that you've learned the names of the animals, and the word game, we'll see how that can help you on the next game."

For testing, Experimenter 2 then took over, and the test screen was placed on the table between Experimenter 2 and the child. Testing was done by putting two of the alien animals (each of a different higher-level category) in front of the child and asking: "Where's the noom/zaag?" The two aliens and the listener stimulus to be presented on each trial were specified by stimulus sheets, which were designed to ensure that Experimenter 2 remained blind to the trained relations and could not cue the child's responses. For the test, four alien animal stimuli were used belonging to two lower-level categories and to two higher-level categories (for example, for Sara two hibs and two tors were used for the first test in Study 1b, the hibs being a kind of zaag, and the tors being a kind of noom). The test consisted of four blocks of 12 trials, with one block per combination of animals. First, both animals from the original Set 1 were presented during one block of trials, followed by both animals from Set 2 in the next block (for Sara this was H1 and T1 for the first block, followed by H2 and T2 for the second block of trials). In the last two blocks of trials the animals were presented in combinations of one animal from Set 1 and one from Set 2 (for Sara this was H1 and T2, and H2 and T1).

Per block of 12 trials, the child was asked six times "where's the noom?" and six times "where's the zaag?". So noom and zaag were each targeted six times per block, three times while in the left position, and three times while in the right position. The trials were presented in a quasi-random order. That is, a random order was generated and then checked and, if necessary, adjusted to avoid obvious regular patterns in the position of the targeted animal (e.g., the correct animal is always on the left). Furthermore, the same trial should not occur twice in a row, where "the same trial" means, for example, noom targeted while the correct animal is in the left position.

The listener behaviour criterion for the higher-level names of the animals was 10 out of 12 correct per alien animal. Using the binomial distribution statistic it was found that 10 out of 12 correct would indicate an above chance performance (to produce 10 or more correct by chance, $p = 0.019$).

No reinforcement was given at this stage, nor was behaviour corrected during trials. However, in the event that a child performed poorly on the first block of trials, Experimenter 1 gave Clue 1 (see Table 2.5). If the poor performance persisted, a second clue was given after Block 2, and a third clue after Block 3.

Table 2.5 Clues given in the event of a child's poor performance.

Clues	
1	Remember the hib-zaag[1] game? What does that tell you? Does that help you to choose the zaag?
2	What goes with zaag?
3	A hib is a zaag!

After testing for listener behaviour at the higher name level with regard to the first two links (e.g., hib-zaag and tor-noom), the intraverbal training procedure was

[1] Particular names differed per child.

conducted for the remaining links (e.g., feb-zaag and lup-noom), followed by another test for listener behaviour at the higher name level.

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 100%. In the same way, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

The word game: echoic training, intraverbal training, and reinforcement reduction. All 8 participants completed the echoic and intraverbal training for the first leg of the hierarchy (see Figure 2.9 for two examples). The number of training trials needed for Leg 1 and 2, for each individual child are shown in Figure 2.10. For Leg 1, on average they needed 98 trials (range 96—104): 33 trials for the echoic training (range 32—40), 33 for the intraverbal training with reinforcement (range 32—40), and 32 for intraverbal training without reinforcement (all children needed 32 trials). The total number of trials to complete all training for females was 96 trials on average (all girls needed 96 trials), while for males it was 99 (range 96—104). Generally, there was little difference in the total number of trials required between participants, and between females and males. Six out of eight participants needed no more than the minimum number of trials to reach criterion in all three stages of the study (32 trials for each stage). Two boys needed one more 8-trial block, in the echoic training stage and the intraverbal training stage with reinforcement, respectively. After completion of training for Leg 1 of the hierarchy, the children were tested for listener behaviour at the higher name level for Leg 1.

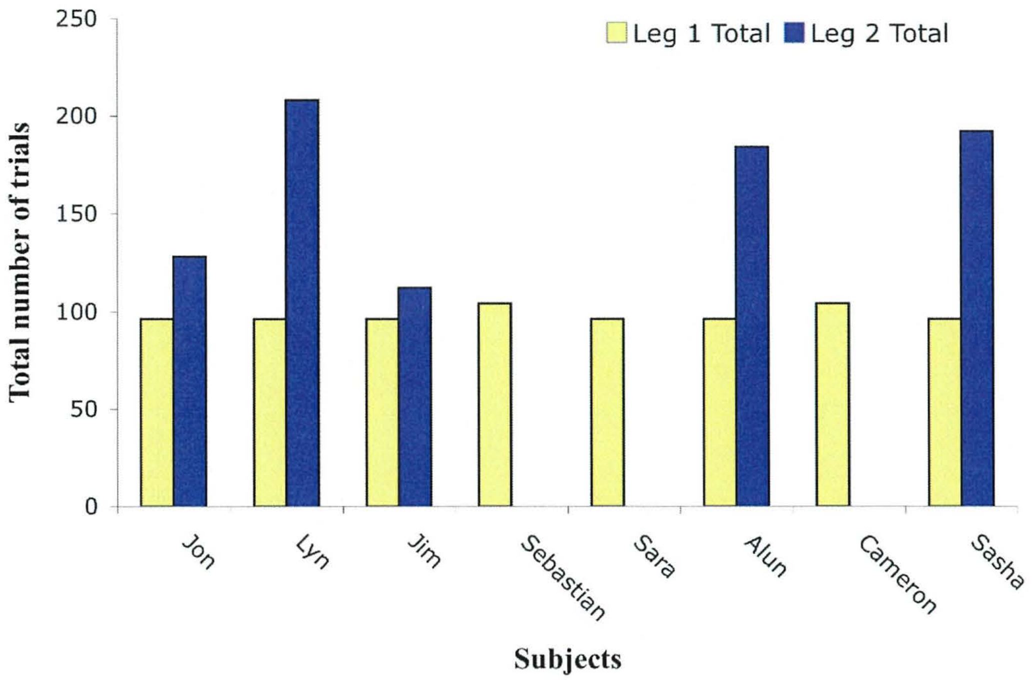


Figure 2.10 The total number of training trials for all individual children in the word game with the alien names (Leg 1 in yellow, Leg 2 in blue).

The word game revisited. Five children started the word game training for Leg 2 of the hierarchy, after being tested for listener behaviour at the higher name level for Leg 1. All five participants completed training. This time, they needed more trials than for Leg 1, on average 165 trials in total (range 112—208): 32 trials for the echoic training (all children needed only the minimum of 32 trials), 99 for the intraverbal training with reinforcement (range 40—144), and 34 for intraverbal training without reinforcement (range 32—40). Females needed on average a total of 200 trials (range 192—208), while males needed 142 (range 112—184). So for females the total number of trials was considerably higher than for males. In intraverbal training with reinforcement all children needed more than the minimum of 32, while in intraverbal training without reinforcement only one boy needed one more 8-trial block.

Listener behaviour test at the higher name level. Figure 2.11 shows the percentage of correct responses per child for each of the zaags and nooms in the test for listener behaviour at the higher name level, for Leg 1 and Leg 2 of the hierarchy. The listener behaviour criterion was 10 out of 12 (83%) correct per animal. Following training for the first two name links, a child was tested for Leg 1 of the hierarchy.

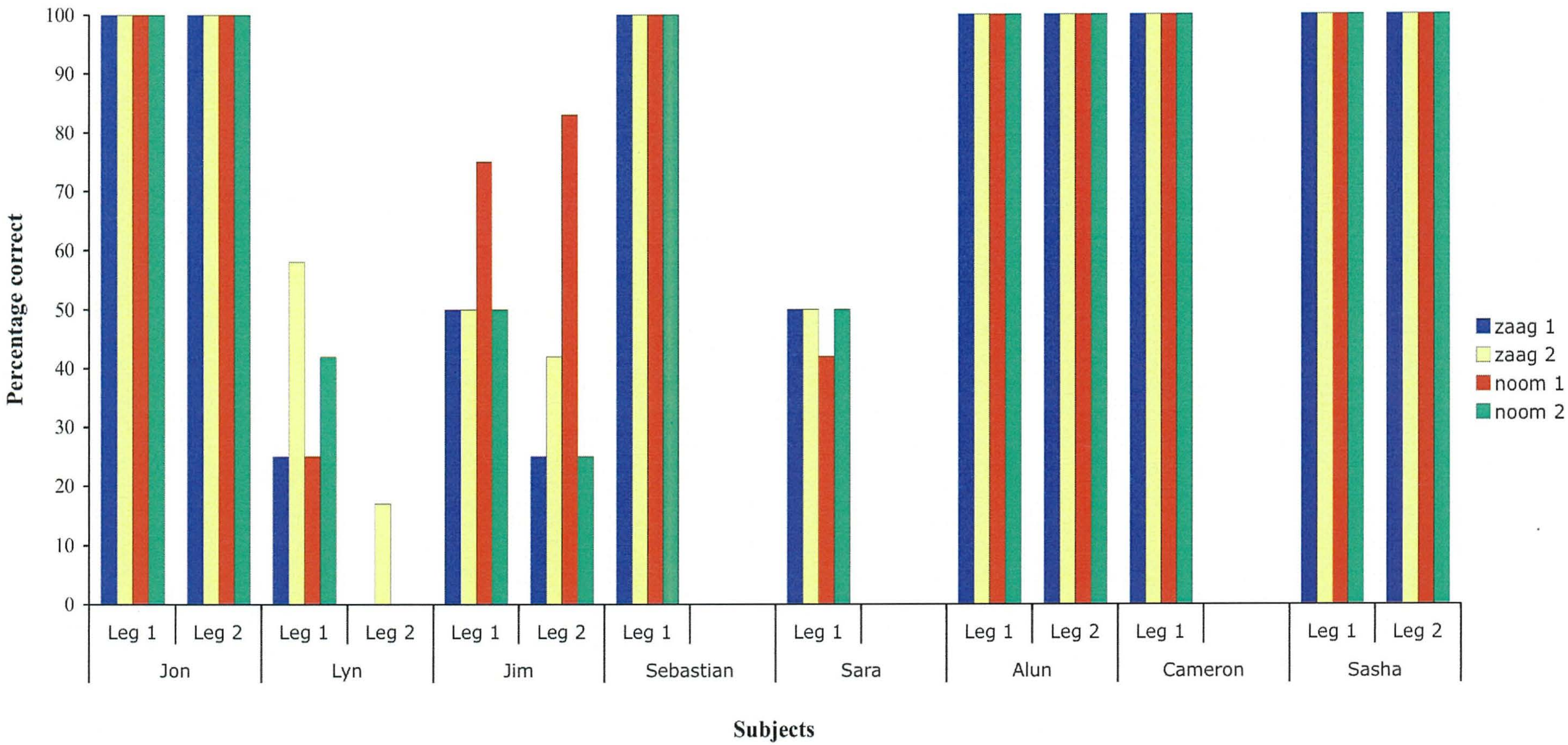


Figure 2.11 The percentage of correct responses per child for all the zaags (in blue and yellow) and nooms (in red and green), separately, in the test for listener behaviour at the higher name level. The listener behaviour criterion was 10 out of 12 (83%) correct per animal. Sebastian, Sara, and Cameron were not available for training and testing for Leg 2.

Leg 1. As can be seen from Figure 2.11, of the 8 participants tested for listener behaviour with regard to Leg 1 of the hierarchy, 5 children passed. They all showed an errorless performance over 48 trials (12 trials per animal). Three children failed the test (Sara, Lyn, and Jim). They scored around chance level (50%) or below. Jim did have 75% correct for Noom 1, but otherwise his performance was at chance level. After the test, Sara, Lyn and Jim were given additional sessions, which are described after the results for Leg 2. The vocalisations by the children during testing for Leg 1 are listed in Table 2.6. After testing, one or more trials were usually repeated, and following a response the child would be asked why that response was given. That is, the child was asked: "Why is that a zaag/noom?" Some children replied that they didn't know, or "cause it is" (Sara), but other children replied by referring to the lower-level name: "cause it's a hib/feb/lup/tor" (Jon, Sebastian, and Alun).

Table 2.6 Children's vocalisations during the listener behaviour test at the higher name level, Leg 1.

		Vocalisations during listener behaviour test	
Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Jon	Feb 1 + Hib 2	Which one is the zaag?	Hib-zaag-hib-zaag
	Feb 2 + Hib 1	Which one is the noom?	That one (points correctly). He's moving. He wants to go.
Cameron	Hib 1 + Tor 1 Hib 2 + Tor 2	Which one is the zaag?	(whispers:) tor-zaag (on 4 trials) tor-zaag
	Hib 1 + Tor 2	Which one is the noom?	noom
		Which one is the zaag?	zaag … tor-zaag
	Hib 2 + Tor 1		zaag zaag is easy the tor-zaag (on 3 trials)
		Which one is the noom?	the noom-tor (on 2 trials)
		Which one is the zaag?	tor-zaag

Leg 2. After the test for Leg 1, intraverbal training for the other two name links took place, and finally listener behaviour testing for Leg 2. Of the 5 participants, 3 produced an errorless performance for this test (as they had done in the test for Leg 1). The 2 children (Lyn and Jim) who had failed Leg 1 also failed Leg 2. The children's vocalisations during testing for Leg 2 are listed in Table 2.7.

Table 2.7 Children's vocalisations during the listener behaviour test at the higher name level, Leg 2.

		Vocalisations during listener behaviour test	
Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Jon	Lup 2 + Tor 2	Which one is the zaag?	(whispers:) tor-zaag-tor-zaag
	Lup 1 + Tor 2	Which one is the zaag?	(whispers:) tor-zaag-tor-zaag
Sasha	Feb 2 + Hib 1	Which one is the zaag?	tor-zaag-tor-zaag. That's the tor, so…

Alun was a special case. When he was tested for Leg 1 his behaviour was errorless. On the day that he was tested for Leg 2, there was a bouncy castle in the nursery and all children were very excited and distracted. Alun unexpectedly showed a complete reversal, scoring 0 out of 12 for two blocks of trials. The test session was discontinued at that point, and it was decided to disregard the session because Alun did not comply with the task. Normally, more training sessions would have been conducted with Alun, but because he would be leaving the nursery within two weeks and it was hoped that he could take part in Study 1c before leaving, an accelerated procedure was employed. Apart from more training trials for the lower-level names and the word game, he was given 48 trials in which two stimuli were presented, one noom and one zaag, as in the listener behaviour test. On each trial he was asked, "Which one is the zaag/noom?" (six trials per alien). When he pointed, either correctly or incorrectly, he was asked, "Why is that the zaag/noom?" He tended to reply "cause it's a hib/feb/tor/lup". Then he was asked, "and what do you say when I say hib/feb/tor/lup?" (in the word game). His reply to this would be followed by "so which one is the zaag/noom?" For his selection of an alien upon this question, he would not receive feedback. He did very well on these trials, and when he was then tested again for Leg 2, his behaviour was errorless (as can be seen in Figure 2.11), as it had been for Leg 1.

Just as for Leg 1 before, Jim and Lyn failed the listener behaviour test for Leg 2, although Jim did reach criterion for Noom 1. Lyn's response pattern was interesting. She showed complete reversal, consistently picking the wrong animal, which led to a zero score for three out of four animals. For the fourth animal (Zaag 2), she scored 17% correct.

Additional sessions. Children who failed the test for either Leg 1 or Leg 2, or both, were given additional sessions. This applied to Sara, Lyn, and Jim.

Sara failed the test for Leg 1, and was given three additional sessions. She had one session in which she was required to echo the explicit rules "A hib is a zaag" and "A tor is a noom" several times. Then she was asked to complete the sentences "A hib is a…" and "A tor is a…" A retest showed no improvement on the listener test. In a second session, in pair wise training Sara was asked for both animals "What is this?", a prompt for the lower-level name. Then, without the animals present, she was asked: "What do you say when I say hib/feb/tor/lup?" (word game trials). Correct responses were reinforced. Then a retest by Experimenter 1 followed, in which feedback was provided in response to the first error. Sara was told which one was the zaag and which one the noom. Her listener behaviour improved for Block 1 (10/12 correct), and Block 2 (9/12 correct, the criterion was 10/12 correct). For Block 3 and 4, the procedure was the same, but *no correction* was given. The child's performance dropped to chance level (6/12 and 6/12 correct). In a third and final session the child was asked to echo "hib-zaag-hib-zaag", and then "a hib is a zaag". This resulted in the child reacting with "a hib is a zaag" to the experimenter's model "hib-zaag-hib-zaag". Praise was provided. The same was done for the tor-noom link.

Following this, with pair wise presentation the child was asked, “Which one is the zaag/noom?” In response to the first error, feedback was given by pointing at the correct animal and asking: “what is this?” When the child answered correctly, the experimenter said, “Remember the word game? hib-zaag-hib-zaag, a hib is a …?” (or: tor-noom-tor-noom, a tor is a …?). Following a correct response, Sara was asked again, “Which one is the zaag/noom?” If the response was incorrect, the prompt procedure was repeated. The child’s listener behaviour improved over four blocks of 12 trials each (4, 8, 9, 8 correct per block), but criterion was still not reached. Interestingly, when the experimenter pointed at the tor and asked “What is this?” the child replied several times: "tor-noom-tor-noom".

Lyn was also given additional sessions after failing the test for Leg 1 and after failing the test for Leg 2. After the first test, she was given one block of test trials with the prompt “which one is the zaag/noom?” for each trial. After each response, she was asked, “why is that the zaag/noom?” Her reply varied from giving one correct and one incorrect lower-level name (on two different trials), to “it’s just the noom”. But mostly she replied, “Don’t know”. She scored 6/12. She was given maintenance training for the lower-level name and the word game, and then provided with a more explicit rule. This was done by having her echo “tor-zaag-tor-zaag … a tor is a zaag” (8 trials) and “hib-noom-hib-noom … a hib is a noom” (8 trials). In the next session the same training was given. Following this, Experimenter 1 conducted one block of test trials, on which Lyn scored 5/12, and finally another block of test trials was given. This time feedback was provided in the following way. After each response, the experimenter said, “why is that a zaag/noom? (…) If I put it on the carousel and ask ‘what’s this?’ what do you say? (the child gave the lower-level name, e.g., tor) And what do you say when I say tor/hib? (…) So which one is the zaag/noom?" This procedure did not lead her to reach criterion; she scored 8/12 correct.

Training for Leg 2 was then started. When she was tested after training for Leg 2, her performance showed complete reversal, as can be seen in Figure 2.11. This meant that for almost all trials (46 out of 48), whenever she was asked for the zaag, she would point at the noom, and vice versa. Lyn was then given more of the verbal prompt sessions along the lines described above. First with maintenance training for the lower-level names and the word game, which were in fact still in place, together with a block of test trials (conducted by Experimenter 1) with the additional question "why?" after each response. In all 12 trials she replied, "Don't know", and she scored 1/12, again a complete reversal. More maintenance training for the lower-level name and the word game followed, and she was provided with the explicit rule for the aliens in Leg 2. This was done again by having her echo "lup-zaag-lup-zaag … a lup is a zaag" (8 trials) and "feb-noom-feb-noom … a feb is a noom" (8 trials). In the next three sessions the same training was given (two more sessions than for Leg 1). Then two blocks of test trials followed, conducted by Experimenter 1. This time feedback was provided as in the additional sessions for Leg 1. That is, after Lyn's responses in the test trials she was asked for the lower-level name of the selected animal, then she was asked for the word game link, and finally the test trial was repeated. Her performance improved in the course of the first trial block, for which she scored 8/12 correct, but not during the second block, for which she scored 2/12 (reversal again). Maintenance training for the lower-level name and the word game was resumed, and in addition she was provided with a new rule for the aliens in Leg 2: Lyn was instructed to echo "lup-zaag-lup-zaag … a lup *is a kind of* zaag" (8 trials) and "feb-noom-feb-noom … a feb *is a kind of* noom" (8 trials). This type of training was given over four sessions, but after the first session the echoing of "lup-zaag-lup-zaag" was followed by the prompt "and then?" and the child was to complete the verbal rule. One block of test trials followed (no feedback, no reinforcement). Lyn had 8/12 correct. After the 12 trials, she was given two more trials, one targeted the zaag (incorrect response), the other the noom (correct response). After both responses she was asked "why?" For the zaag, she said, "cause it's got horns", for the noom "I don't know". After maintenance training for the lower-level names, the word game, and the new verbal rule, in the next session two more blocks of test trials were conducted (no feedback, no reinforcement). Lyn scored 3/12 and 0/12. When asked which one the zaag was and why, she selected incorrectly, and said again "cause it got horns". A final session, before she left the nursery, was like the previous session but with one block of test trials. Lyn scored 0/12 (complete reversal again).

Finally, there was Jim who also failed the test for Leg 1 and was given three special sessions. In all three he was given maintenance training for tacting the aliens at the lower name level, and for the word game, on which he did very well overall, except that in tacting in the first session he made one error with regard to Feb 2, and two for Hib 2. In subsequent trials his responses were correct for these aliens. Apart from that, in the first session he was retested on two blocks of test trials by Experimenter 1. There was no cueing, no feedback (correction) and no reinforcement, but after each response the child was asked, "Why is that a zaag/noom?" On each trial the child replied, "cause it's a feb/lup". Jim had 9/12 and 10/12 correct, and in the five trials where he pointed to the wrong animal, he did tact those aliens correctly at the lower-level name. So when he was asked for the noom (Feb 1 or 2) and pointed to the zaag (Lup 1 or 2) instead, when asked why, he said, "cause it's a lup". In a second session, this procedure was repeated with a third block of test trials. Jim had 9/12 correct, but interestingly, when he picked the wrong alien as zaag or noom, he also provided the incorrect lower-level names. The three trials with incorrect selection were then repeated, and he now had 2/3 correct. In a third and final session, this procedure was repeated with a fourth block of test trials. Jim had only 2/12 correct. This time he was given feedback in the following way. When after 4 trials he again provided the incorrect lower-level name ("feb" for Lup 1), he was asked, "Is that the feb?" He corrected his response, and when asked again, "which one is the noom?" he pointed correctly and provided the correct lower-level name. In trials in which he was asked for the noom, for example, but selected the zaag and correctly called it lup, he was then asked, "how does the word game go again? What do you say when I say lup?" In all cases, he provided the correct link. Then he was asked, "so which one is the noom?" and in all cases he then selected correctly. In the course of the trial block his performance did not improve because of this feedback, as his score of 2/12 makes clear. Instead of more such sessions Jim was then given word game training for Leg 2 of the hierarchy, and tested for listener behaviour at the higher level at the end of training. He scored 5, 5, 7 and 4 for the four 12-trial blocks, as can be seen in Figure 2.11. Before he left the nursery, he was given one more special session consisting of one 12-trial block with feedback as in the last special session for Leg 1. His score was 5/12.

Discussion

The aim of Study 1b was to investigate how children learn higher-level names for objects and entities. Do they have to be explicitly told the higher-level names in the presence of the objects or entities, or is it perhaps enough for them to just hear and say the lower-level and higher-level names in sequence? Each of the lower-level names hib, feb, tor, and lup were linked intraverbally with a potential higher-level name, either zaag or noom, in an echoic and intraverbal word game that was played without the aliens being present. Once the word game was learned, the children were tested for listener behaviour with regard to the potential higher-level names (i.e., hear /higher-level name/ – select animal).

The data show that of the 8 children who completed training and were tested, 5 passed the listener behaviour test at the higher name level for Leg 1, and 3 of the 5 children who passed the test for Leg 1, also passed the test for Leg 2 (the other 2 children who had passed the test for Leg 1 had left the nursery). So for 5 out of 8 children, learning the intraverbal relation between lower and potential higher-level names was enough for them to also acquire listener behaviour at the higher name level, without direct training. This is in line with Horne and Lowe's description of interchange of listener behaviour across intraverbal names: "when names are reliably linked within an intraverbal sequence, listener behaviour of any one of the name relations will be increasingly evoked by the other name relations and vice versa" (1996, p. 210). Interestingly, of the 5 children who passed the test, in one or more trials 3 children (Jon, Cameron and Sasha) produced the intraverbal name relations out loud before responding in the listener test, while the 2 other children (Sebastian and Alun), along with Jon, referred to the name links in explaining their responses ("Why is it a noom?" – "Cause it is a feb/tor/hib/lup").

With regard to the children who failed the test, the data show that extra sessions, in which they were given more explicit rules in a further word game, did not help them to pass the test at later stages. So even after such sessions, Lyn and

Sara's behaviour during the listener behaviour test at the higher name level had not come under the control of the intraverbal relations they had learned, or the explicit rules. When during the test Sara was given one of the clues as listed in Table 2.5, and was thus reminded of the word game, she even asked, "Is that important?" suggesting this lack of control by the intraverbal relations. For Jim, the sessions in which he was asked to explain why he had selected a particular alien, initially seemed to improve his performance, but in further sessions his performance deteriorated.

Study 1c – Names and transfer of behaviour

Imagine a child playing with a small plastic bowl (an example described by Horne & Lowe, 1996, p. 213). When the child is told during play that it is a boat, she may then show behaviour appropriate to boats. For example, she may put it in water, or pretend water, and have it make floating movements. Alternatively, if the child is told it is a hat, she may put the container on her head and show other behaviours associated with hats. In these cases the listener behaviour (or functional responding) with regard to boats and hats, transfers to the plastic container by means of the names. This is called *transfer of function* (see, for example, Lowe, et al., 2005, and Horne, et al., 2006). Once a child has learned names at different levels for particular objects, she may then also show transfer of function at different name levels. This name-based transfer is the focus of Study 1c.

In this study, two functional responses were trained (i.e., see animal – produce gesture), one to a zaag, one to a noom. The functional responses can be seen as greetings, and are particular movements of the hand, in one case, left hand touching left shoulder, and in the second case hands touching each other at the fingertips while palms are facing down (see Figure 12). Following training, the children were tested for transfer of function to the other stimuli. Would the trained greetings transfer only at the lower name level or also at the higher name level? If for one particular child, one greeting was trained to Hib 1 (a zaag) and the other greeting to Lup 2 (a noom), then would the greeting learned for Hib 1 only transfer to Hib 2, or also to the other zaags (let's say, the two febs), and would the greeting learned for Lup 2 only transfer to Lup 1, or also to other nooms (in this case the two tors)?[2] This was tested by presenting the child with the aliens not used in function training, one zaag and one noom at a time, and asking the question "How does this one go?" while pointing at one of the two aliens (i.e., see animal – produce gesture).

Method

Participants

Of the participants in Study 1b who passed the test for listener behaviour at the higher level for all legs of the hierarchical structure (see Figure 2.9), 3 children (1 female and 2 males) were available to take part in Study 1c. They were between the ages of 3 years and 11 months, and 5 years and 2 months at the start of the study (see Table 2.8).

Table 2.8 Participants' sex, age at start of training, and age at first listener behaviour test.

Participant	Sex	Age at start (years/months*)	Age at testing (years/months*)
Jon	M	3/11	4/0
Alun	M	4/1	4/1
Sasha	F	5/2	5/3

* age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).

Apparatus and Stimuli

Apparatus and stimuli in Study 1c, and all the following studies, were as in Study 1a.

[2] Bear in mind that intraverbal name links varied for individual children. So a hib could be a zaag for one child, while it was a noom for another child.

Procedure

As before, a single case design with replication across participants was employed. No parents were present during the study. There were one or two daily sessions with the child, which took about 10 to 20 min each.

Arbitrary stimulus tact training and intraverbal training. For each participant, Study 1c was started as soon as possible after completion of Study 1b. Where necessary, the tact relations of Study 1a and the intraverbals of Study 1b were retrained to criterion.

Novel function training. For each child, two of the eight alien animal stimuli, one zaag and one noom, were randomly selected for novel function training. The two novel responses to be trained, were 1) hand on shoulder, and 2) hands in front, touching each other at the fingertips while palms are facing down (see Figure 2.12).

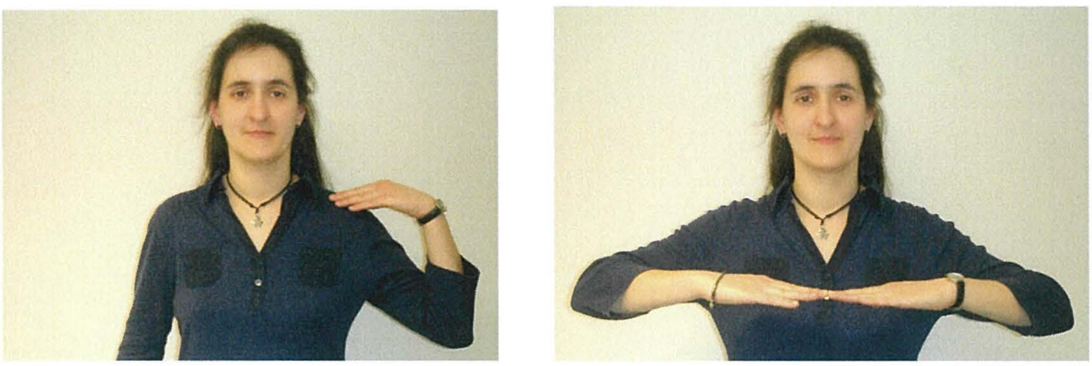


Figure 2.12 The functional responses each trained to a different alien stimulus.

Each response was randomly assigned to a particular animal for each child. The relations between the specific animals, their names at both levels, and the trained functional responses for each participant are shown in Table 2.9. During training the two alien animal stimuli were placed in front of the child. In the first trial for each alien, the experimenter pointed at the target stimulus and introduced it by saying:

"Look, this one goes like this [experimenter demonstrates gesture]. Can you do this?" In subsequent trials with the same alien, the experimenter pointed at the targeted stimulus and said: "How does this one go?" Following a correct response the experimenter delivered social praise (saying e.g., "Yes, well done!")[3]. If the child did not produce the correct response, corrective feedback was provided: "It goes like this. Can you do this?" One stimulus was targeted on each trial. The aliens were each targeted four times (each twice on the left, and twice on the right) in a pre-specified quasi-random order in blocks of eight trials such that the same trial type (e.g., hib on left while targeted) did not appear twice in succession. The order of targeting was listed on a stimulus sheet. The learning criterion for this stage was 7 out of 8 correct over two consecutive 8-trial blocks. During the function-training period, maintenance training was conducted with regard to tacting of the aliens (Set 1 and 2), and with regard to the word game. The trained and tested relations in Study 1c are shown in Figure 1.4 (in Chapter 1).

---

[3] And occasionally, in addition, a music book was presented on which the child could press a button.

Table 2.9 For each child, the novel behaviour trained to the specified stimulus (gestures are shown in Figure 2.12).

Participant	Trained Gesture	
	Hand on shoulder	Hands in front
Alun	 Feb 2 / zaag	 Hib 1 / noom
Sasha	 Lup 1 / noom	 Hib 2 / zaag
Jon	 Tor 1 / zaag	 Lup 2 / noom

Reduction in reinforcement rate. As before, in the last training stage reinforcement was reduced to 0%. This was done to make sure that the child's performance remained at criterion in the absence of reinforcement, in preparation for the test situation in which there would be no contingent reinforcement delivered. In these no reinforcement sessions the experimenter used the penguin hand puppet (Peter Penguin), and asked the child to teach Peter the game. If the performance of any child dropped below criterion, then behaviour was to be trained to criterion again.

Category transfer-of-function test. Following novel function training, the child was tested for transfer (i.e., see animal – produce gesture) of these responses to the other aliens. Prior to testing, Experimenter 1 conducted review trials for tacting of the aliens in Set 1 and 2. Each of the eight stimuli was targeted once, to determine whether the trained tact behaviour was still in place. Review trials were also conducted for the word game (four trials per link, e.g., hib-zaag), and for function training (four trials for each of the two animals). Furthermore, it was checked whether listener behaviour for the higher-level names was still in place (four trials for each of the four randomly selected pairs of one zaag and one noom). After these trials, Experimenter 1 sat behind the child, while Experimenter 2 sat opposite the child.

The test screen was placed between Experimenter 2 and the child, to control for cueing. Of the six alien animals not used in function training, two animals (a random selection of one zaag and one noom) were put on the table in pre-specified positions. Experimenter 2 pointed at one of the two animals, and asked the child: "How does this one go?" If the child did not respond within 4 s, the verbal prompt was repeated up to twice more if necessary. Each stimulus was targeted ten times in total (five times on the left and five times on the right) over two 10-trial blocks. The stimuli were targeted in a pre-specified quasi-random order in which the same trial (e.g., Feb 1 on left and targeted) did not occur twice in succession. The order of targeting was listed on stimulus sheets on which the names and shapes were not linked, to control for cueing.

In this manner, all six stimuli were targeted ten times, making a total of 60 trials. The criterion was 80% (8 out of 10) correct for each alien, over two consecutive blocks. An additional criterion for behaviour transfer was 80% (8 out of 10) correct responses emitted to each animal pair, over two consecutive blocks. Using the binomial distribution statistic it was found that 8 out of 10 correct per animal pair would indicate a performance well above chance; the probability of producing 8 or more per animal pair correct by chance is 0.044. When this criterion is applied over two consecutive blocks, this probability is $<.001$. The combined criterion was necessary, because it was possible to reach criterion for an alien by producing the same gesture on many or all trials in two 10-trial blocks, where it would be doubtful whether this would be a case of transfer of function. A child who reached criterion for one alien in a particular pair in this way, did not reach criterion for that pair. No reinforcement was given during the test.

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 100%. In addition, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Function training and reinforcement reduction. All 3 participants completed the function training. The average number of trials to criterion was 64 (range 40—88); 43 trials for the function training with reinforcement (range 24—56), and 21 for function training without reinforcement (range 16—32). The total number of trials to criterion for both training stages for the female participant was 40 trials, while for males it was 64 and 88, respectively. After reaching criterion for the function training, the children were given more training trials, because their response latencies were large (an estimated 4 to 6 s; sometimes it was necessary to repeat the prompt). Trials of overtraining were added for strengthening the responding and decreasing latency, that is, until the child responded immediately and correctly on X consecutive trials. Jon had eight extra trials, while Alun had 16, and Sasha 40.

Transfer of function test. Table 2.10 shows the alien pairs used in training and testing, and the number of correct responses out of 10 per pair for each child. The transfer of function criterion was 80% (8 out of 10 responses) correct for each of the two 10-trial blocks per alien pair. Only Jon reached this criterion for all three pairs. Alun reached it for Pair 1, and Sasha reached criterion for none of the pairs. The other criterion was 80% percent (8 out of 10) correct for each alien. Figure 2.13 shows the percentage of correct responses per child to each of the six aliens presented in the test (i.e., the aliens not involved in training). Two functions had been trained, one to a zaag (e.g., Hib 1), one to a noom (e.g., Feb 2). Figure 2.13 shows whether the functions transferred without direct training to the other members of the lower-level categories (in this example, Hib 2 and Feb 1), and whether they transferred to the other members of the higher-level categories (the other zaags and nooms, in this case, the tors and the lups).

Table 2.10 The alien pairs used in training and testing, and the test scores for the three alien pairs (P1, P2, P3), for each child. The transfer of function criterion was 80% (8 out of 10) correct for each of the two 10-trial blocks per alien pair. The top score for each pair shows the result for the first test trial block, the second score gives the result for the second trial block.

Child	Gesture	Training stimuli	Test stimuli: Pair 1	Test stimuli: Pair 2	Test stimuli: Pair 3	Scores per pair (out of 10)
	Hand on shoulder					P1: 9 10
Alun		Feb 2/zaag	Lup 1/zaag	Lup 2/zaag	Feb 1/zaag	P2: 7 1
	Hands in front					P3: 4 6
		Hib 1/noom	Hib 2/noom	Tor 2/noom	Tor 1/noom	
	Hand on shoulder					P1: 4 5
Sasha		Lup 1/noom	Feb 1/noom	Lup 2/noom	Feb 2/noom	P2: 4 5
	Hands in front					P3: 5 5
		Hib 2/zaag	Tor 2/zaag	Tor 1/zaag	Hib 1/zaag	
	Hand on shoulder					P1: 10 10
Jon		Tor 1/zaag	Tor 2/zaag	Hib 2/zaag	Hib 1/zaag	P2: 10 10
	Hands in front					P3: 10 9
		Lup 2/noom	Feb 1/noom	Lup 1/noom	Feb 2/noom	

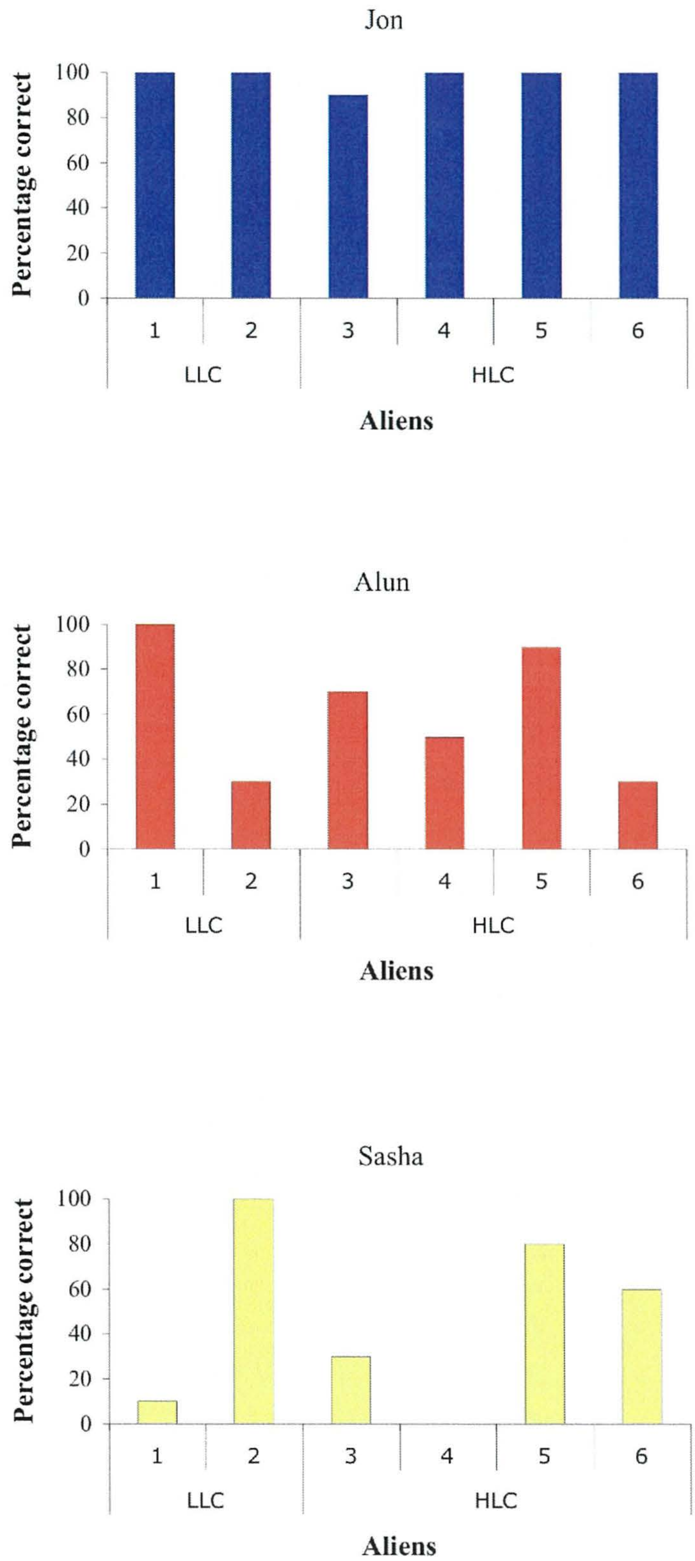


Figure 2.13 The percentage of correct responses for all six aliens used in transfer of function testing, two of the same lower level category (LLC), and four of the same higher level category (HLC) as the aliens used in function training. Criterion was 80% correct per alien.

Jon showed almost perfect transfer of function both at the lower level and the higher level. He made just one error in 60 trials (there were 10 trials for each of the six aliens). Alun showed transfer of function to two aliens, and Sasha seemed to do so as well; in both cases it concerned transfer to an alien belonging to the same lower-level category as one of the aliens used in training, and to an alien belonging to the same higher-level category as one of the aliens involved in function training. However, Sasha reached criterion for these aliens by producing the same gesture on almost every trial, over two blocks.

The vocalisations produced by the children during testing are listed in Table 2.11. Sasha's vocalisation was produced in the initial test, not in retesting. As in the previous study, one or more trials were repeated in a post-test interview. The experimenter asked the child: "Why does it go like that? Jon replied again by referring to the lower-level name: "'Cause it's a feb". Experimenter 1 then asked: "And is that how febs go?" Jon confirmed this.

Table 2.11 The children's vocalisations during transfer of function testing.

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Alun	Lup 2 + Tor 2	How does this one go?	Looks like that (compares shape of hands with shape of alien, while making gesture right next to alien: incorrectly produces the hands in front gesture next to Lup 2. It is unclear to the experimenter why Alun sees similarities in these shapes. In a later trial he again suggests a similarity in shape of gesture and alien, when he correctly makes the hands in front gesture, a "long gesture", one could say, next to Tor 2, a long shape. He first did this in training, producing the hands in front gesture right next to Feb 2, again a long shape. All three shapes are shown below*).
Sasha	Lup 2 + Tor 1		tor-zaag (then produces correct gesture)

* Alun's aliens:

Lup 2  Tor 2  Feb 2 

Additional sessions. Sasha was given additional sessions after the test. These were comparable to the initial training sessions for the functional responses, except that the verbal prompt was more elaborate. Instead of being asked, "How does this one go?" for the two aliens involved in function training, she was asked, "What's this? (...) Is it a zaag or a noom? (...) And how does it go?" Along with this, she received maintenance training for the lower-level names and the word game. This extended function training was given over four 8-trial blocks with reinforcement, and two blocks without reinforcement. Sasha was then tested by Experimenter 2, using the same elaborated prompt on every trial. Over Block 1, 2 and 3 her performance was errorless: her answers to the first part of the verbal prompt were correct on all trials, and for the second part of the verbal prompt she only made an error on one trial, but she still produced the correct gesture for that alien. In Block 4 she had 8/10 correct, she always gave the correct lower-level name, and made only one error for the second part of the prompt (in the two trials in which she produced the incorrect response, her answers to the first and second part of the prompt were correct). Unfortunately, the test session was unavoidably long, and her performance regressed over the final two blocks: 6/10 (two errors for the second part of the prompt, but not in the trials in which she produced the incorrect gesture) and 4/10 (five errors for the same prompt, but in three of the six trials in which she produced the incorrect gesture, she replied correctly both the first and second part of the prompt). For these last two blocks she went back to producing mainly the same gesture (hands in front) over two trial blocks. Sasha was then given six more training blocks with the extended verbal prompt, and was then tested again for the last two trial blocks of the earlier test. Her performance was not better than before (2/10 and 6/10 — she made one error in response to the first part of the prompt, and 11 in response to the second part)). After two more training blocks, she was eventually tested one last time on the same two blocks, and she scored 5/10 and 8/10 (making four errors in response to the first part of the prompt, and four in response to the second part). The percentage of correct responses for Feb 2 (Alien 4) was 10% and increased to 60%, while for Hib 1 (Alien 2 — see Figure 2.14 below) it was 90% and decreased to 70%.

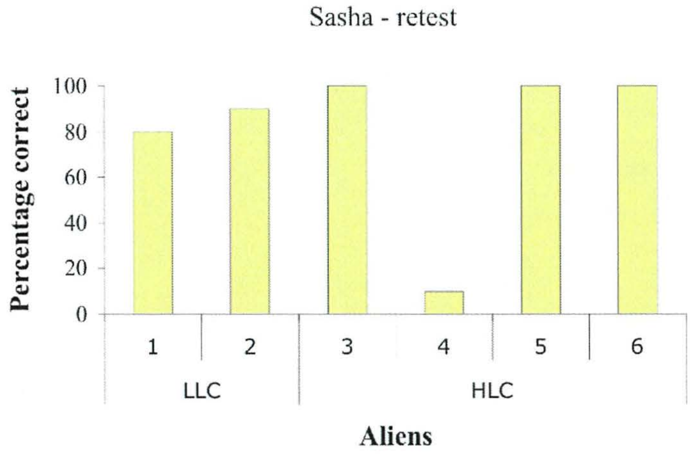


Figure 2.14 The retest data for Sasha. The percentage of correct responses for all six aliens used in transfer of function testing, two of the same lower level category (LLC), and four of the same higher level category (HLC) as the aliens used in function training. The additional transfer of function criterion was 80% (8 out of 10) correct for each alien.

Alun initially appeared to be very confused about what was expected of him. He was therefore given this instruction: "It can go like this or like this (Experimenter 1 demonstrates both gestures: hands in front, and hand on shoulder). Just pick either of the two". Although he did respond in the trials that followed, he was very hesitant in his responses. In addition to that, the test session was unavoidably long, and he got tired as the session progressed. Whereas his performance was good for Pair 1 (9/10 for Block 1, and 10/10 for Block 2), it deteriorated from then on; he had low scores for Pair 2 (7/10 and 1/10) and Pair 3 (4/10 and 6/10). When the percentage of correct responses per alien are considered, he showed transfer of function for both aliens of Pair 1, but for none of the others.

It was suspected that Alun might be able to perform better if retested with Pair 2 and 3. Attempts were made to provide a better situation in the retest than in the initial test. For example, the review trials before testing (in which tacting at the lower level was checked in a few trials, along with the word game, the listener behaviour at the higher level, and the trained functional responses) were done in a separate session. Also, the retest was done over two short sessions (on the same day), one session for Pair 2, and one for Pair 3.

At the start of both of the short retest sessions an attempt was made to encourage Alun to respond without worrying about getting it right or not. However, this had to be done without reinforcing any transfer of function behaviour directly. Therefore, he was first given two trials for each of two aliens not used in function training (Hib 1 and Feb 2), without reinforcement. He responded correctly on all trials. Then he was given two trials for each of the two aliens that were involved in function training (Hib 2 and Lup 1). When he responded correctly on all trials for these aliens, Experimenter 1 said: "that was very good!"

After these trials, the retest trials started. For Pair 2, he now scored 4/10 (Block 1) and 6/10 (Block 2). For the individual aliens, he showed transfer of function to Tor 2, but not for Lup 2. For Pair 3, he scored 4/10 (Block 1) and 5/10 (Block 2). He showed transfer of function to neither of the two aliens. So in all, he showed transfer of function to one more animal in the retest, bringing the total to three aliens, one of the same lower level category as the aliens involved in training, and two of the same higher level category as the aliens in training.

Discussion

The aim of Study 1c was to investigate name-based transfer of function. Would responses (gestures, comparable to greetings) taught to one member of one higher-level category (one zaag) and one member of the other higher-level category (one noom) transfer to the aliens that were not involved in training? And if so, would there be transfer just at the lower name level, or also at the higher name level?

The data show full transfer of function at both levels for only 1 of the 3 participants, Jon. Alun reached criterion for one pair of aliens, while Sasha reached criterion for none of the pairs. Their percentage correct for the individual aliens did suggest that both Alun and Sasha showed transfer of function for two aliens, one of the same lower-level category and one of the same higher-level category as one of the aliens involved in training. But Sasha achieved this by almost always producing one gesture in all trials over two trial blocks, which gives her transfer of function a questionable status. Her vocalisation (see Table 2.11) indicates that during the initial test she was making use of the word game. The trial in which she said "tor-zaag" out loud was one of four trials on which she responded correctly.

Alun's data are interesting because he showed transfer of function at both levels, but he did not show consistent transfer of function to either of the two levels. It could be the case that his transfer at the higher-name level was brought about through transfer of function at the lower name level, in combination with "selection by exclusion". That is, because he was trained to produce the functional responses to Feb 2 and Hib 1, when presented with Lup 1 and Hib 2 he could infer the correct gesture for Hib 2, and then on the basis of that he could infer that the other alien in this pair (Lup 1) would require the production of the other gesture. But why does his behaviour not transfer to Feb 2, the other alien in the same lower-level category as the aliens involved in training? Test sessions were relatively long (due to review trials and the large number of test trials needed), and perhaps this influenced the children's performance as well. Alun showed his best performance in the first two trial blocks. On a different note, in one training trial and two test trials he placed his hands next to an alien while making one of the two gestures, and he commented that they (the shape of the hands and the shape of the alien) look the same. This indicates that at least in some trials he may have based his responses on shape rather than name. That said, his responses are not consistent, because he first made the hands in front gesture next to Lup 2, and later next to Tor 2. Jon's response to the why-question after testing shows that he did base his behaviour on (at least the lower level) names.

In general, transfer of function as shown by Jon and Alun is in line with Horne and Lowe's (1996, p. 213) description of how functional responding may transfer by means of names (see the example given in the introduction for Study 1c, earlier in this chapter). The children in these studies have learned names at different levels for particular objects, and now they show (full or partial) transfer of functions at different name levels. But did the children actually learn "names" at both levels? They did indeed learn to *name* at the lower level: they reached criterion in tact training and they passed the listener behaviour test. However, at the higher name level, they showed appropriate listener behaviour, after echoic and intraverbal training in the word game (Study 1b), but it is unclear whether the children had learned speaker behaviour at the higher name level, because they had not received any direct training for this speaker behaviour, and they had not been tested for it either. So some children may have been able to tact the aliens at the higher name level, while others may not have been able to do this, and it is possible that this influenced their performance in the transfer of function test.

After additional sessions with the more elaborate verbal prompt, when required to name the stimuli at both levels before producing the gesture, Sasha reached criterion for the first two pairs, but not for the last pair. Her percentage of correct responses for the individual aliens show that there was now transfer of function for five out of the six aliens, two of the same lower level category and three of the same higher level category as one of the aliens involved in training. This is a substantial improvement compared to the earlier test, due to the experimenter prompting the child to use the names at the lower as well as the higher level in both training and testing.

Study 1d – Names and transfer of a second behaviour

To extend transfer of function data and to see whether transfer of behaviour would be consistent within the same child — that is, whether a child showed a regular behaviour pattern in testing — a study similar to the previous one was conducted. The only difference between Study 1c and 1d was that the trained novel response was not a gesture (comparable to a greeting), but an alien animal cry. Two cries were chosen: "Boo!" and "Raagh!" These cries were trained to two animals, one to a zaag and one to a noom. After training the child was once again tested for transfer of function to the other stimuli in the respective higher-level categories (see animal – produce cry).

Method

Participants

Of the participants in Study 1c only 1 male participant (Jon) was available to take part in Study 1d. He was 4 years and 2 months old at the start of the study, and 4 years and 3 months old at testing.

Procedure

Apart from the functional responses, the procedure was exactly the same as in Study 1c. Therefore a brief summary will suffice here.

Arbitrary stimulus tact training and intraverbal training. Because some time had passed since the completion of Study 1c, before starting Study 1d the tact relations of Study 1a and the intraverbals of Study 1b were retrained to criterion.

Novel function training. Again, two of the eight alien animal stimuli, one zaag and one noom, were selected for this second round of novel function training. These two stimuli belonged to different lower-level categories than the ones used for function training with this participant in Study 1c (a tor and a lup were used for training in Study 1c, whereas a hib and a feb were employed in Study 1d). The two novel responses to be trained were animal cries. The child was told that one animal said "Boo!" and the other said "Raagh!" Each response was randomly assigned to one animal. The relations between the specific animals, their names at both levels, and the trained functional responses are shown in Table 2.12. In the first trial for each alien, the experimenter pointed at the target stimulus and introduced it by saying, "Look, this one says (experimenter produces the relevant cry). Can you say that?" In subsequent trials with the same alien, the experimenter pointed at the targeted stimulus and said: "What does this one say?" Following a correct response the experimenter delivered social praise[4]. Targeting and order of stimuli were the same as in Study 1c, as was the learning criterion. Maintenance training during the function-training period was conducted not only with regard to tacting of the aliens and the word game, but also for the functions trained in Study 1c. The trained and tested relations in Study 1d are shown in Figure 1.5 (in Chapter 1).

---

[4] And occasionally, in addition, a music book was presented on which the child could press a button.

Table 2.12 The novel behaviours, the cries "Boo!" and "Raagh!" were allocated to one zaag and one noom.

Participant	Cry	
	Boo!	Raagh!
Jon	 Hib 1 / zaag	 Feb 2 / noom

Reduction in reinforcement rate. As before, in the last training stage, reinforcement was reduced to 0%. The experimenter used the penguin hand puppet (Peter Penguin), and as in previous test phases, asked the child to teach Peter the game.

Category transfer-of-function test. The child was tested for transfer of these vocal responses to the aliens that were not involved in training (see animal – produce cry). As in Study 1c, review trials preceded testing (tact trials with Set 1 and 2, word game trials, and function training). Then Experimenter 2 tested the child for transfer of the vocal responses to the aliens not used in training, by pointing at one of the two animals on the table (always one zaag and one noom), and asking the child: "What does this one say?" Presentation and targeting of the stimuli, as well as the number of trials, and the criterion for behaviour transfer was exactly the same as for Study 1c. No reinforcement was given during the test.

Interobserver reliability. All trials in a randomly selected 25% of all training sessions were scored by an independent observer; interobserver agreement on these trials was 100%. Similarly, all test trials were scored; interobserver agreement on these trials was 100%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

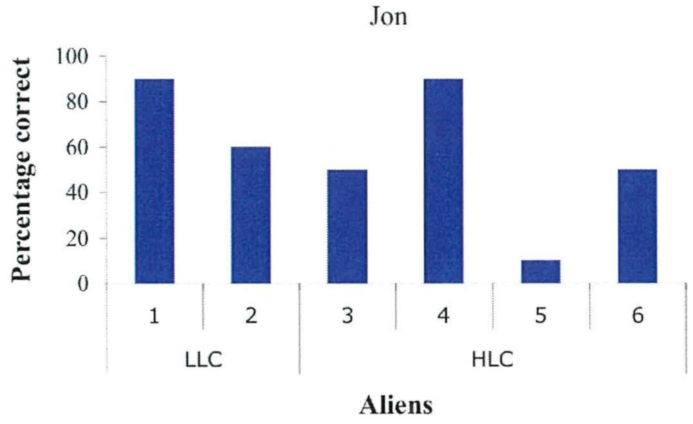
Function training and reinforcement reduction. Jon reached criterion for novel behaviour training in 40 trials (which was the minimum), 24 trials for the training with reinforcement, and 16 for training without reinforcement.

Transfer of function test. Table 2.13 shows the alien pairs used in training and testing, and the scores per pair. The transfer of function criterion was 80% (8 out of 10) correct for each of the two trial blocks per alien pair. Jon only reached this criterion for Pair 2. The additional criterion was 80% (8 out of 10) correct responses per block, over two consecutive blocks, for each alien. Figure 2.15 (top graph) shows the percentage of correct responses to each of the six aliens not involved in training. As in Study 1c, two functions had been trained, one to a zaag (for Jon this was Hib 1), one to a noom (for Jon Feb 2). Figure 2.15 shows whether any untrained transfer of function was found within the lower-level categories (in this case, transfer to the category members Hib 2 and Feb 1), or within the higher-level categories (that is, to the other zaags and nooms, in this case the tors and the lups).

Table 2.13 The alien pairs used for Jon, in training, and in testing and retesting, and the scores for the three alien pairs (P1, P2, P3) in the test, and the retest. The transfer of function criterion was 80% (8 out of 10) correct for each of the two 10-trial blocks per alien pair. The top score for each pair shows the result for the first test trial block, the second score gives the result for the second trial block.

Cry	Training stimuli	Test stimuli: Pair 1	Test stimuli: Pair 2	Test stimuli: Pair 3	Scores per pair (out of 10): Test	Scores per pair (out of 10): Retest[5]
Boo					P1: 10 0	P1: 7 3
	Hib 1/zaag	Tor 1/zaag	Tor 2/zaag	Hib 2/zaag	P2: 9 9	P2: 8 10
Raagh						
	Feb 2/noom	Lup 2/noom	Feb 1/noom	Lup 1/noom	P3: 5 2	P3: 10 10

[5] In the retest, the pairs were presented in a different order: P3, P1, P2. For comparability, the retest scores are given in the same order as the scores for the initial test.



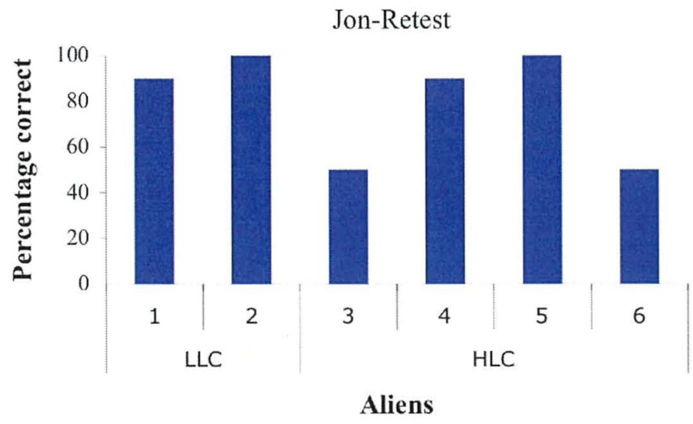


Figure 2.15 The percentage of correct responses for all six aliens used in transfer of function testing, two of the same lower level category (LLC), and four of the same higher level category (HLC) as the aliens used in function training. The additional transfer of function criterion was 80% (8 out of 10) correct per alien.

Jon's performance did not reach the additional criterion for all animals. Transfer of function to two aliens was found; to an alien belonging to the same lower-level category as one of the aliens used in training, and to an alien belonging to the same higher-level category as one of the aliens involved in function training. So the almost perfect transfer of function in Study 1c, was not repeated in Study 1d.

Because of the difference in results between the two studies, Jon was given 48 extra training trials (in which he never scored below criterion, which was 7/8 per block). Then he was retested, and the results can be found in the bottom graph of Figure 2.15.

This time he did better, although not as well as in Study 1c. He showed transfer of function to two more animals. All in all, the trained functional responses had now transferred to both aliens in the same lower-level category as the aliens used in training, and to two out of the four aliens that belong to the same higher-level category.

Jon's vocalisations during testing are listed in Table 2.14. As before, one or more trials were repeated in a post-test interview. Following a response the experimenter asked the child, "Why does it say that?" This time Jon wasn't sure, and instead of referring to the lower-level name (e.g., "cause it's a feb"), he produced other responses: "Cause it's a boo-animal", "Because it likes that", "Because he lost his boo-voice" (for an animal that says Raagh!), and "Because he lost his raagh-voice" (for an animal that says Boo!).

Table 2.14 Jon's vocalisations during testing for transfer of function (Test 1d).

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Jon	Tor 1 + Lup 2	What does this one say?	It does this (hand on shoulder-gesture), but it says Boo!
	Tor 1 + Lup 2		lup-noom-lup-noom
	Feb 1 + Tor 2		It's a feb and it says Raagh!
	Hib 2 + Lup 1		It's a hib but it says ... I don't remember. Think it says Boo!
	Hib 2 + Lup 1		It's a hib and it says Raagh!
	Hib 2 + Lup 1		(with hesitation:) Boo! (E2: Are you sure?) Yes, he really says Boo!
	Retest:		
	Lup 1 + Hib 2		Raagh, because I know the hib says Boo!
	Tor 1 + Lup 2		Boo! A tor with a lup? A feb with a hib.
	Feb 1 + Tor 2		I forgot. (E1: think about it.) Raagh!
	Feb 1 + Tor 2		Boo! Because that one (Feb 1) says Raagh!

The underlined stimuli are the targeted stimuli.

Discussion

As in Study 1c, the aim of this study was to see whether functional responses (animal cries) taught to one zaag and one noom, would transfer to other aliens, within the lower level category and also within the higher level category.

The data show partial transfer of function at both levels, in the initial test as well as the retest. Jon reached criterion only for Pair 2, in the initial test, and for Pair 2 and 3 in the retest. The percentage correct for the individual aliens showed transfer of function for two aliens initially, one of the same lower level category and one of the same higher level category as one of the aliens involved in training. In the retest, Jon showed transfer of function to two more aliens, again one of the same lower level category and one of the same higher level category as one of the aliens involved in training. So initially Jon showed transfer of function at both levels, but like Alun in Study 1c he did not show consistent transfer of function to either of the two levels.

However, in the retest, Jon did show consistent transfer at the lower level. The two aliens he did not show transfer to were members of the same higher-level category as the aliens involved in training. However, these two aliens were presented together. So the two cases of transfer at the higher level could have been brought about through transfer of function at the lower name level, in combination with "selection by exclusion" (as explained with regard to Alun's results in Study 1c). But it should also be noted that in Study 1c Jon showed transfer of function to two aliens belonging to the same higher-level categories as the ones involved in training. If he did not need to rely on selection by exclusion there, then he may not have needed to do so in the present study either. His vocalisations during the retest, however, suggest that he may have responded by exclusion.

In general, Jon's (partial) transfer is in line with Horne and Lowe (1996 – see the previous section). His vocalisations indicate that his behaviour is again based on names.

Study 1e – Names, transfer of behaviour, verbal prompts and stimulus control

Study 1e was a test only, focussing on the effects of different verbal prompts on the child's transfer of functional responses. It investigated the effects of more general prompts than those used in Study 1c ("How does this one go?") and 1d ("What does this one say?"). So for example, "What does this one *do*?" or "What *can* this one *do*?" Would the more general verbal prompts have stimulus control over both types of functional responses (the gestures and cries taught in the previous studies) in the absence of direct training? In other words, would the new prompt evoke both types of responses without direct training?

In addition to this, the test aimed to extend the transfer of function data even further. In Study 1c and 1d the child was trained in two different contexts to respond with a total of four separate functional responses to four alien animal stimuli, each belonging to different lower-level categories. Would these responses (see animal – produce gesture/cry) transfer to the other stimuli in the lower-level categories, as well as to the other stimuli in the higher-level categories?

Method

Participants

One male participant (Jon) was available to take part in Study 1e. He was 4 years and 9 months old at testing.

Procedure

The procedure was as in the preceding studies, unless stated otherwise below.

Arbitrary stimulus tact training and intraverbal training. Where necessary, the tact relations of Study 1a, the intraverbals of Study 1b, and the functional behaviour trained in Studies 1c and 1d were retrained to criterion.

Category transfer-of-function test. The child was tested for transfer of all four previously learned functional responses to all eight aliens (see animal – produce gesture and cry). As before, review trials preceded testing (8 alien animal tact trials, 16 word game trials, 4 trials per link, and 4 trials for each of the 4 functional responses). Then Experimenter 2 tested the child for transfer of previously learned functions to all aliens, by pointing at one of the two animals on the table (always one zaag and one noom), only providing a different verbal prompt than before: "What *can* this one *do*?"[6] Apart from the verbal prompt, the test only differed from the previous tests in the total number of trials. In Study 1c and 1d the participants were tested for transfer of function to the six stimuli not involved in training, whereas this time all eight stimuli were presented. As before they were presented in pairs of one zaag and one noom. . Each alien animal stimulus was targeted 10 times, making a total of 80 trials. Targeting of the stimuli, and the criterion for behaviour transfer were exactly the same as for Study 1c and d. No reinforcement was given during the test.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 100%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

---

[6] Initially a different verbal prompt was used: "What does this one *do*?" However, the test session using this prompt had to be ended after only one test sheet, because the participant was not at all sure what was expected of him. The prompt was then changed and the trials on that test sheet repeated.

Results

Transfer of function test. Study 1e was a further test for transfer of function that also focussed on the verbal prompt and stimulus control. In Study 1c the child was prompted to produce gestures (hand on shoulder, and hands in front, see Figure 2.12) with the question “How does this one *go*?” In Study 1d the prompt for the production of animal cries (“Boo!” and “Raagh!”) was the question “What does this one *say*?” Table 2.15 shows the animals these behaviours were trained to in the two studies, for Jon. In Study 1e the verbal prompt was more general than in the previous two tests: “What does this one *do*?” Table 2.16 lists the order of presentation of the alien pairs.

Table 2.15 Overview of the functional responses and the animals they were trained to.

	Gesture		Cry	
Participant	Hand on shoulder	Hands in front	Boo!	Raagh!
Jon	 Tor 1/zaag	 Lup 2/noom	 Hib 1/zaag	 Feb 2/noom

Table 2.16 The alien pairs in Test 1e. After Session 1 the verbal prompt was changed.

Session	Alien pair*	
1	Hib 2	Lup 2^
	Tor 2	Feb 1^

2	Hib 2	Lup 2^
	Tor 2	Feb 1
	Tor 1	Feb 2
	Hib 1	Lup 1
3	Tor 1	Feb 2
	Hib 1	Lup 1
	Tor 2	Feb 1
4	Hib 2	Lup 2
	Hib 2	Lup 2

*One 10-trial block per pair. ^Blocks repeated at the end.

Four sessions were needed to complete Test 1e. In Session 1, in the first trial block, Jon produced animal cries only, to both aliens (4/10 correct). In the second trial block, Jon appeared to be confused about what was expected of him. In Trial 1 and 2 (Feb 1 targeted in both), he replied "feb-noom". When in Trial 3 (Tor 2 targeted) he repeatedly said, "This is too hard", the session was discontinued.

The next day the test was resumed, with a slightly changed verbal prompt. The prompt "What *does* this one *do*?" was changed to "What *can* this one *do*?" In Trial 1 (Lup 2 targeted) with the new prompt, Jon responded with a (correct) gesture. In Trial 2 (Hib 2 targeted) he was confused again, and said, "I really don't know". The rest of the trial block was then left for later. Instead, Jon was tested on the next block, in which he consistently produced gestures for both Tor 2 and Feb 1 (10/10 correct). He did the same in the next two trial blocks, with Tor 1 and Feb 2 (10/10 correct), and Hib 1 and Lup 1 (10/10 correct).

A few days later, a further two blocks of test trials were conducted. In these blocks, Jon consistently produced the correct gesture for Tor 1 and the correct animal cry for Feb 2 (10/10 correct), and then the correct cry for Hib 1 and the correct gesture for Lup 1 (10/10 correct). So he now produced both types of functional responses within trial blocks, but not within one trial. That is, he did not produce both responses with regard to one alien.

Finally, Jon was retested on the two trial blocks that had not been completed, and on the first trial block (used at the start of the test with the initial prompt). In these trial blocks he resumed his "strategy" of producing a gesture for one alien, and a cry for the other: he produced the gesture for Tor 2 and the cry for Feb 1 (10/10 correct), and in the last two trial blocks he produced the gesture for Lup 2 and the cry for Hib 2 (10/10 + 10/10 correct). So eventually, the more general verbal prompt had control over both types of functional responses (gestures and cries), but the type of response that was produced, depended on the particular animal that was targeted.

In sum, when only taking into account the trial blocks with the second verbal prompt ("What *can* this one *do*?"), the results show that initially Jon produced only gestures for three pairs of animals, and all these gestures were correct. By doing this, he showed perfect transfer of function at the lower name level, for Tor 2 and Lup 1 (Jon was trained to produce the gestures to Tor 1 and Lup 2). He also showed perfect transfer of function at the higher name level, for Hib 1, Feb 1 and Feb 2. In the next session he switched to producing correct gestures to some animals, and correct cries to the others. Here, he produced the correct gestures to the aliens that were involved in gesture training (Tor 1 and Lup 2) and correct cries to the aliens involved in training for the cries (Hib 1 and Feb 2), and he showed perfect transfer of function at the lower name level only: he produced the correct gestures to Tor 2 and Lup 1, and cries to Hib 2 and Feb 1. Jon's vocalisations during testing are listed in Table 2.17.

Table 2.17 Jon's vocalisations during testing.

		Vocalisations during Test 1e	
Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Jon	Hib 2 + Lup 2	What does this one do?	It boos.
	Hib 2 + Lup 2		It roars.
	Hib 2 + Lup 2		Talks too much.
	Feb 1 + Tor 2		feb-noom
	Feb 1 + Tor 2		This is too hard! (Trial repeated) It eats pasta shapes.
	Hib 2 + Lup 2	What *can* this one do?	I really don't know.
	Tor 1 + Feb 2		The tor has been here and the feb. I just don't know.
	Tor 2 + Feb 1		Feb says raagh.
	Tor 2 + Feb 1		It can turn into anything.
	Tor 2 + Feb 1		Don't know. (E1: you knew before, or you thought you knew before.) Raagh.

The underlined stimuli are the targeted stimuli.

Discussion

The first aim of Study 1e was to see whether a more general verbal prompt than the ones used in Study 1c and d[7], namely "What can this one do?" would evoke both of the previously learned functional responses (gestures and cries) in the absence of direct training.

The data show that the more general prompt did indeed evoke both functional responses. However, Jon would produce both types of functional responses within

[7] The verbal prompt in Study 1c was "How does this one *go*?", and in Study 1d it was "What does this one *say*?".

trial blocks, but not within one trial, so not with regard to one alien. The type of response that was produced (gestures or cries) depended on the particular animal that was targeted. Interestingly, Jon produced just gestures for the first three pairs of animals, while for the other pairs (in a different session) he switched to producing gestures to some animals, and cries to the others. More specifically, he produced gestures to the aliens that were involved in gesture training and to the aliens with the same lower-level name as these, and cries to the aliens involved in training for the cries and to the other aliens with the same lower-level name. It is not clear what made Jon switch to a different "strategy" of responding. There were a few days between the two test sessions. If the test session had not been broken in separate parts, there may not have been such a switch. It is clear from Jon's vocalisations during the test, that he was initially struggling with the task.

Study 1e was also a further test for transfer of both functions. All functional responses Jon produced were correct: he showed perfect transfer of function at the lower and the higher name level for the gestures, for the first three pairs of animals. In the cases of transfer of function at the higher name level, the second pair contained Tor 1, the alien involved in Jon's gesture training, so higher level transfer could have been brought about through transfer of function at the lower name level, in combination with "selection by exclusion" (as explained before).

In the case of the alien pairs for which he produced gestures to some animals and cries to others, he showed perfect transfer of function at the lower level.

In general, Jon's (partial) transfer is in line with Horne and Lowe (1996), as explained above.

Study 1f – Names, functional responses and listener behaviour

In Study 1c, two functional responses (gestures, comparable to greetings — see Figure 2.12) were trained, one to a zaag, one to a noom. Children were taught to produce the correct functional response when asked about a targeted alien animal: "How does this one go?" (see animal – produce gesture). Study 1f is a test only. Instead of focussing on transfer of the above named functions, this study investigates whether without being directly trained to do so a child would demonstrated the corresponding listener behaviour with regard to the functional responses (i.e., see gesture – select animal), when asked "Which one goes like this (E2 demonstrates gesture)?"

Method

Participants

One male participant (Jon) was available to take part in Study 1f. He was 4 years and 9 months old at testing.

Procedure

The procedure was also the same as in the preceding studies, unless stated otherwise below.

Listener behaviour test. The child was tested for listener behaviour (i.e., see gesture – select animal) with regard to the functional responses trained in Study 1c. Prior to testing, Experimenter 1 conducted review trials for the trained functional responses (2 trials for each of the two animals).

For the test, only the two stimuli (one zaag and one noom) to which the child was trained to emit the novel responses in Study 1c were placed on the table, in pre-specified positions. During test trials it was not possible to use the test screen, because the experimenter needed to produce the gestures and make sure they were clearly visible to the child. All efforts were made to avoid cueing the child's behaviour. Note that Experimenter 2, who tested the child, did not know which gestures were appropriate for which aliens.

Experimenter 2 tested the child's listener behaviour by pointing at one of the two animals and asking the child: "Which one goes like this (E2 produces gesture)?" If the child did not respond within 4 s, the verbal prompt was repeated up to two more times. Each of the two stimuli was targeted ten times in total (5 times on the left and 5 times on the right) over two 10-trial blocks. The stimuli were targeted in a pre-specified quasi-random order in which the same trial (e.g., Tor 1 on left and targeted) did not occur twice in succession. The order of targeting was listed on stimulus sheets on which the names, the shapes, and the functional responses were not linked, to further reduce possible cueing. With both stimuli targeted ten times, there was a total of 20 trials.

The listener behaviour criterion for the functional responses was 8 out of 10 correct per animal pair, over two consecutive blocks, and per alien. Using the binomial distribution statistic it was found that 8 out of 10 per block correct would indicate a performance well above chance; the probability of producing 8 or more per block correct by chance is 0.044. When this criterion is applied over two consecutive blocks, the probability that this score is due to chance is <.001. No reinforcement was given during the test.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 100%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Listener behaviour test for functional responses. Study 1f tested Jon for listener behaviour with regard to the two functional responses (the gestures) the child had learned to produce to two aliens in Study 1c (Tor 1 and Lup 2). Jon showed almost perfect performance, scoring 9/10 correct in the first block, and 10/10 in the second. The listener behaviour criterion was 80% (8 out of 10) correct per block, over two consecutive blocks, and 80% (8 out of 10) correct per alien. For Alien 1, Jon produced correct responses on 90% of the trials, and for Alien 2 he had 100% correct. Jon produced no vocalisations during testing.

Discussion

The aim of Study 1f was to investigate whether children who had learned to produce two functional responses (gestures) with regard to one zaag and one noom, would demonstrate the corresponding listener behaviour upon seeing the gesture (i.e., see gesture – select animal), without being directly trained to do so.

The data show that listener behaviour with regard to the functional responses was indeed in place when Jon had learned to produce the gestures to the aliens. This could be seen as parallel to the emergence of listener behaviour in the course of vocal tact training, which was conducted in Study 1a, and is another example that some behaviour can be established without the need for direct reinforcement/training.

Study 1g – Names, functional responses and listener behaviour, once again

In Study 1d, another two novel responses (the alien animal cries "Boo!" and "Raagh!") were trained, one to a zaag, one to a noom. The child was taught to produce the correct vocal response when asked about a targeted alien animal: "What does this one say?" (see animal – produce "cry"). Study 1g is a test only. As in the previous study, it investigates whether, without being directly trained to do so, a child would demonstrate the corresponding listener behaviour with regard to the vocal responses (i.e., hear /cry/ – select animal).

Method

Participants

One male participant (Jon) was available to take part in Study 1g. He was 4 years and 9 months old at testing.

Procedure

Apart from the particular novel responses, the procedure was exactly the same as the procedure in Study 1f, so only a brief summary will be given here.

Listener behaviour test. The child was tested for listener behaviour (i.e., hear cry – select animal) with regard to the novel vocal responses trained in Study 1d. Prior to testing, Experimenter 1 conducted review trials for the trained novel vocal responses (two trials for each of the two animals). For the test, only the two stimuli (one zaag and one noom) employed in novel vocal response training in Study 1d were placed on the table, in pre-specified positions. Experimenter 2 tested the child's listener behaviour by pointing at one of the two animals and asking the child, "Which one says (E2 produces cry)?" Each of the two stimuli was targeted ten times in total (five times on the left and five times on the right) over two 10-trial blocks, so overall there were 20 trials. Presentation and targeting of the stimuli, as well as the number of trials, and the criterion for listener behaviour to the novel vocal responses was exactly as in Study 1f. No reinforcement was given during the test.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 100%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Listener behaviour test for functional responses. Again, Jon was the only participant. In Study 1g, he was tested for listener behaviour to the two vocal responses (the animal cries "Boo!" and "Raagh!") that he had learned to produce to two aliens (Feb 2 and Hib 1) in Study 1d. Jon's performance was errorless over 20 trials (10 per animal), thereby reaching both the listener behaviour criterion per block (80%; 8 out of 10 over two blocks) and the additional criterion of 80% (8 out of 10) correct per alien. Jon produced no vocalisations during testing.

Discussion

As for Study 1f, the aim of Study 1g was to investigate whether children who had learned to produce two novel responses (animal cries), one to one of the zaags and another to one of the nooms, would demonstrate the corresponding listener behaviour upon hearing one or other of these animal cries (i.e., hear /cry/ – select animal), without being directly trained to do so.

As in Study 1f, the data show that listener behaviour to the novel vocal responses was indeed in place when Jon had learned to produce the cries to the aliens. This could be seen as parallel to the emergence of listener behaviour in the course of vocal tact training, which was the subject of Study 1a, and is another example that not all forms of behaviour need direct reinforcement/training to be brought about.

Study 1h – Category match-to-sample test

Horne and Lowe (1996) proposed that categorization may be brought about without direct training, by learning the same name to a set of arbitrary stimuli (i.e., stimuli that share no characteristics apart from the name; so they have no physical features such as shape or colour in common, that distinguish them from other stimuli). In Study 1h this was tested by presenting the arbitrary alien animal stimuli, which the child had learned to name earlier, singling out one of the stimuli and asking the child to give "the others" (see one category member – select the others). Could the child, without any direct categorization training, sort the stimuli into categories on the basis of the common name? And would a child show category sorting only at the lower name level, or also at the higher name level?

Method

Participants

One male participant (Jon) was available to take part in Study 1h. He was 4 years and 9 months old at testing.

Procedure

The procedure was also as in the preceding studies, unless stated otherwise below.

Category match-to-sample Test 1. The child was tested for category matching-to-sample (i.e., see one category member – select the others). Review trials, conducted by Experimenter 1, preceded testing (eight alien animal tact trials, and 16 word game trials, four trials per link, and for checking listener behaviour with regard to the higher-level names there were two trials for each of the four randomly selected pairs of one zaag and one noom).

After the review trials, all eight alien animal stimuli were placed on the table in random pre-specified positions. Experimenter 2 conducted the category matching-to-sample test by holding up one of the eight animals as the sample stimulus and asking the child, "Look at this one; can you give me the others like this one?" If the child did not respond within 4 s, the verbal prompt was repeated. The child was to select the other animals belonging to the same higher-level category. So when the experimenter held up a noom (let's say Lup 1), the child was to pick up the other three nooms (not just Lup 2), or the response was classed as incorrect. When the child responded by handing over one or more animals, Experimenter 2 asked, "Any more?" (to avoid cueing the child, this was asked after any selection, not just when the child had not handed over enough animals). If the child said "no" to this, the stimuli were removed and repositioned for the next trial.

Each of the eight stimuli served as the sample six times (48 trials). The sample stimuli were selected in a pre-specified quasi-random order, which ensured that the same sample stimulus did not occur twice in succession. The order of targeting was listed on stimulus sheets on which the names at both levels and the shapes were not linked, to prevent cueing by the experimenter. The category match-to-sample criterion for the lower and the higher level together was 3 out of 6 correct category sorts per alien animal as sample. Using the binomial distribution statistic it was found that 3 out of 6 correct would indicate a performance well above chance; the probability of producing 3 or more correct by chance is <0.001. For categorisation at the lower level, the criterion was correct selection of the other lower level category member on 16 out of 48 trials ($p<.001$). No reinforcement was given during the test.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 98%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Category match-to-sample test. Jon passed the category match-to-sample test with a correct selection of aliens in 43 out of 48 trials. In these 43 trials he not only showed correct category sorting at the lower name level, but also at the higher name level. It can be seen from Figure 2.16 that for five of the aliens Jon made one wrong selection per sample animal, while for the other three aliens as samples he made no errors. The criterion was 50% (3 out of 6) correct sorts per sample animal. In the five trials in which Jon made a wrong selection, he always picked up the correct lower-level category member. His errors either involved picking up too many animals, or picking up one or two wrong animals. Table 2.18 lists Jon's vocalisations during the test.

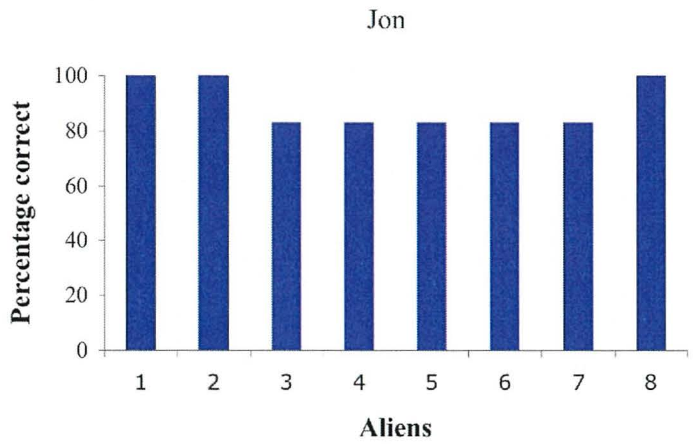


Figure 2.16 The percentage of correct selections for each of the eight aliens in the category match-to-sample test. The criterion was 50% (3 out of 6) correct per alien.

Table 2.18 Jon's vocalisations during match-to-sample testing.

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Jon	Hib 1 (zaag) +all others	Look at this one. Can you give me the others like this one?	(First trial. Jon picks up Hib 2. E2: Are there any more?)
			I can't see any hibs.
			(E2 repeats prompt, then Jon picks up the other zaags; Tor 1 and Tor 2.)
	Tor 2 (zaag) +all others		(After making a correct selection:) They're all zaags.

Discussion

The aim of Study 1h was to investigate whether a child could sort the alien animal stimuli into categories solely on the basis of the common name (see one category member – select the others), that is, without any direct categorization training. If correct sorting did occur, would the child show category sorting only at the lower name level, or also at the higher name level? Jon showed correct category sorting not only at the lower name level, but also at the higher name level.

This is in line with Horne and Lowe (1996) who proposed that categorization may be brought about without direct training, by learning the same name to a set of arbitrary stimuli (i.e., stimuli that share no characteristics apart from the name; so they have no physical features in common, like shape or colour, that set them aside from other stimuli). In this case, the children had learned the same lower-level names for pairs of stimuli (e.g., Hib 1 and 2), and these lower-level names were related intraverbally to the potential name at the higher level. The same higher-level names applied to combinations of four stimuli (e.g., Hib 1 and Hib 2, and Feb 1 and Feb 2 could all be zaags). Jon's comment after a correct selection of higher-level category members ("They're all zaags") suggests that he did indeed base his selections on the names.

* * *

In the second series of studies (2a-h), described in the next chapter, all the alien animal studies described so far were repeated. However, this time pre-training with familiar shapes was added for every study. That is, before starting the procedure with the aliens or alien names for each study, the same procedure was employed with stimuli that were toy versions of real-life shapes (foods and animals). As in the alien studies, the procedure involved training and testing. The added pre-training was designed to ensure that the children understood the test instructions, before the arbitrary stimuli were introduced. Would the children need a smaller number of trials to criterion for the tasks involving aliens, because the procedure was familiar from pre-training? And would more children pass the tests, having seen the models of real-life objects embedded in procedures analogous to those employed subsequently in the training and testing tasks with the aliens?

Chapter 3 - Study 2

Study 2a – Vocal tact training, preceded by pre-training with familiar stimuli

Study 2a, like Study 1a, focussed on learning of name relations, and on the relation between tact behaviour and the corresponding listener behaviour. It tested the naming account in the same way as Study 1a (training of tact behaviour for the same eight alien animals, followed by testing for listener behaviour), the only difference being that pre-training with *familiar stimuli* preceded the procedure with the *aliens*. For this pre-training, apart from the stimuli that were used, the procedure was also exactly the same as that in Study 1a; it consisted of tact training (i.e., see stimulus– say "name") with eight stimuli belonging to four different categories, followed by testing for the corresponding listener behaviour (i.e., hear /name/ – select shape).

The aim of Study 2a was to see whether children would learn to tact the aliens faster (i.e., whether they would need a smaller number of trials to criterion) when they were first exposed to all the training and testing procedures with familiar stimuli, using the same instructions as were subsequently given during training and testing with the unfamiliar stimuli—the alien animals. The aim was to ensure that the children fully comprehended the instructions so that this would not be a limiting factor in the children's performances with the alien stimuli (see Horne, et al., 2006, p. 268).

Method

Participants

Participants were 13 children (4 females and 9 males) between the ages of 3 years and 3 months, and 4 years 4 months at the start of the study (see Table 3.1). They were recruited from Tir na n-Og, the Daycare Nursery and Centre for Child Development at the University of Wales Bangor. Initially 3 more children were recruited (2 females, 1 male), but unfortunately they were not willing to participate any further after one or two sessions. As noted in Chapter 3, according to routine assessments conducted by the nursery nurses, the participants showed no sign of developmental delay. And again, for all studies, recruitment of the children was subject to parental consent, and ethics approval was gained from the School of Psychology Research Ethics Committee.

Table 3.1 Participants' sex, age at start of pre-training with familiar stimuli, and age at start of tact training, and at first listener behaviour test, with aliens.

Participant	Sex	Age at start of pre-training (years/months*)	Age at start of alien tact training (years/months*)	Age at listener testing with aliens (years/months*)
Tomos**	M	3/3	3/4	3/4
Ginny	F	3/4	3/5	3/8
Huw	M	3/4	3/4	3/5
Glyn	M	3/5	3/7	-
Kevin	M	3/5	3/7	3/8
Leona	F	3/5	3/6	3/8
Adam	M	3/6	3/8	3/9
Cecelia	F	3/6	3/7	3/7
Louise	F	3/8	3/9	3/10
Jamie	M	3/9	3/10	3/11
Mike	M	3/10	3/10	3/11
Simon	M	4/1	4/2	4/4
Lee	M	4/4	4/6	4/9

* age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).

** For reasons of confidentiality, these are not the children's real names.

Apparatus and Stimuli

In both parts of Study 2a (pre-training, and the procedure with the aliens), the experimental setting was exactly as in Study 1a, apart from the different stimuli being used in pre-training. The experimental stimuli used in pre-training were eight familiar stimuli, belonging to four different categories: a snake and a crocodile (both reptiles), a shark and a stingray (both fish), an apple and a banana (both fruits), and a mushroom and a carrot (both vegetables). The stimuli were made out of yellow Fimo and were approximately 5 x 7.5 cm in size (see Figure 3.1). As in Study 1a, the stimuli were presented on carousels.

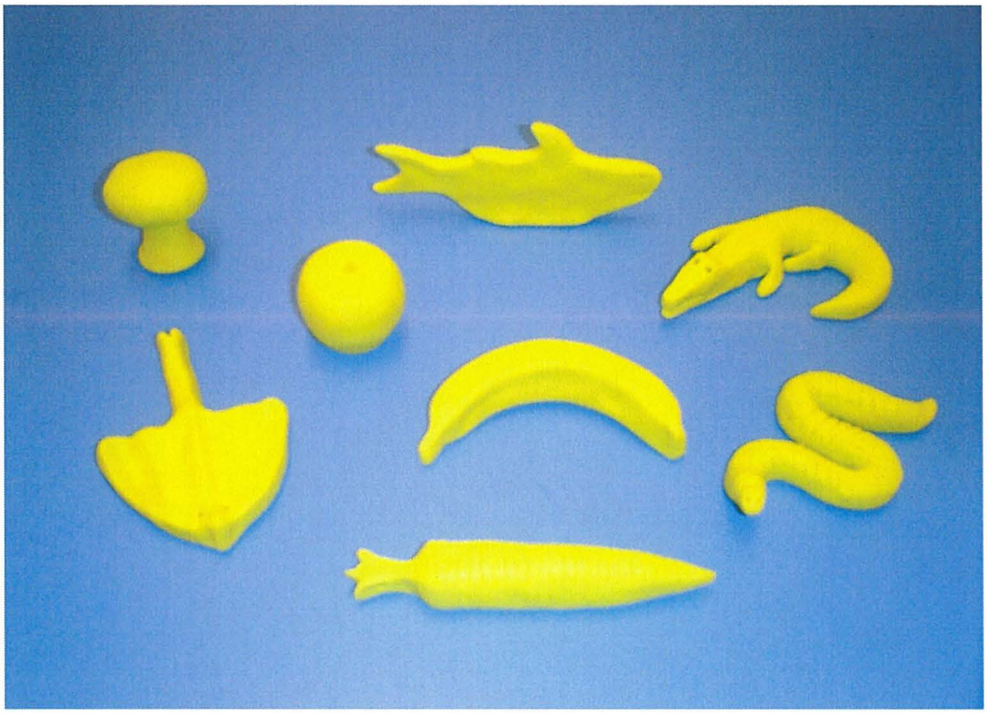


Figure 3.1. The familiar stimuli made out of Fimo

Stimulus sheets were used to schedule the presentation of both the familiar stimuli in pre-training and the alien animals in the second part of the study. To ensure that Experimenter 2 could not cue the child's responses during listener behaviour test sessions, the stimulus sheets employed specified the positions in which the aliens should be placed on each trial and the listener stimulus to be presented: Experimenter 2 was therefore unable to determine the name trained for each alien. In all listener behaviour test sessions, both with familiar stimuli and with aliens, the wooden test screen was used (see Study 1a).

During training sessions, social praise was provided as reinforcement. In addition to this, sometimes a music book was used for reinforcement: the child could press a button for each, or every three, correct responses during the session. At the end of each session the child selected a story to be read by the experimenter. Near the end of all experimental work, magic cards were introduced as reinforcers. These were silver or gold cards, in different sizes, on which a child could put stickers received at the end of a session. Once the magic card was filled up with stickers, the child could take it home.

Procedure

A single case design with replication across participants was employed. Throughout the study there were no parents present. At all stages in the study there were one or two daily sessions with the child, varying in duration from 10 to 20 min.

Tact training—Familiar stimuli. First, the children were familiarised with Experimenter 1 during play in the common nursery rooms and the experimental room. When a child was comfortable in both settings, the child was taken into the experimental room and pre-training with the familiar stimuli was started. For each child the eight familiar stimuli (see Figure 3.1) were randomly assigned to two sets of four stimuli, so that each set contained one reptile, one snake, one fruit and one vegetable.

Apart from the stimuli, all procedures were exactly as in Study 1a. On the first trial for each familiar shape, the experimenter pointed at the target stimulus and introduced it by saying, "Look at this, it's a reptile/fish/fruit/vegetable. Can you say reptile/fish/fruit/vegetable?" On every subsequent trial with the same shape, the experimenter pointed at the target stimulus and said, "What's this?" Correct responses were followed by social praise (saying e.g., "Yes, well done!"), and incorrect responses by corrective feedback: "It's a reptile/fish/fruit/vegetable. Can you say reptile/fish/fruit/vegetable?" The child was trained to criterion first with Set 1 (Reptile 1, Fish 1, Fruit 1, Vegetable 1), then with Set 2 (Reptile 2, Fish 2, Fruit 2, Vegetable 2). During training for Set 2, maintenance trials were conducted with Set 1. After the learning criterion was reached for both sets, the stimuli were systematically interchanged in the mixed sets training and testing.

Mixed sets. The reptiles (R), the fish (Fi), the fruits (Fr) or the vegetables (V) in both sets were exchanged, one category at a time. For example, first the fish were exchanged, then the vegetables, then the reptiles, then the fruits. The order of interchanging was randomised for each child.

Reduction in reinforcement rate. In preparation for the test, randomly generated mixed sets consisting of two stimuli from Set 1 and two from Set 2 were presented (the mixed sets consisted of one stimulus from each category – e.g., R2, Fi1, Fr1, V2), in a 0% reinforcement stage. As in the first study, during no reinforcement sessions the experimenter introduced the penguin hand puppet (Peter Penguin), and asked the child to teach Peter the names of the shapes. At this stage, as before, the learning criterion was. If the performance of any child did not meet the criterion of seven out of eight correct within a block over two consecutive blocks, then reinforcement was reintroduced until responding was once again at criterion.

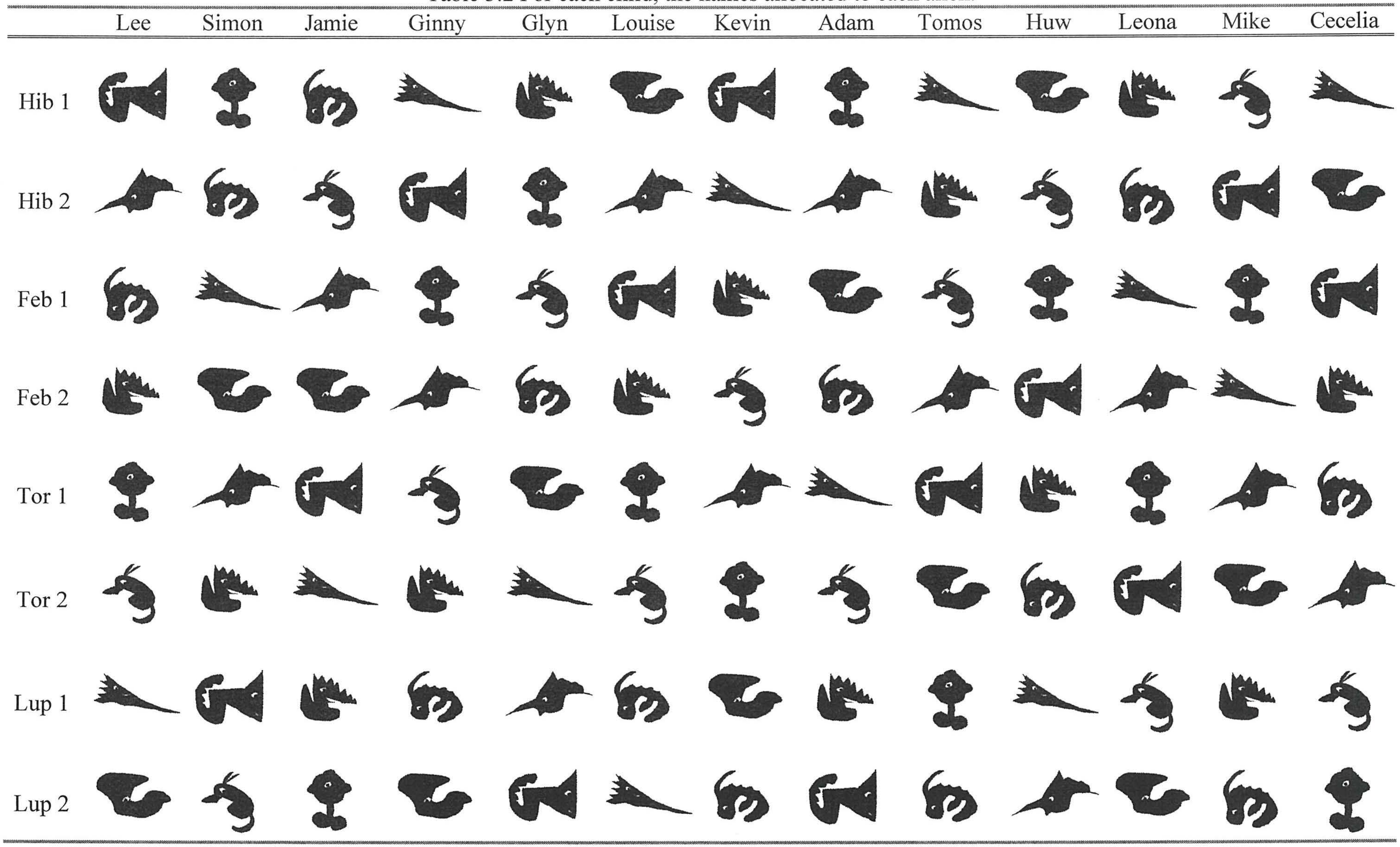
Testing under extinction resumed, and so on, until the child's performance met the zero reinforcement criterion.

Listener behaviour test—familiar stimuli. After tact training, the child was tested for listener behaviour (hear name – select shape) discriminations with regard to the familiar stimuli. At this stage, as in the previous stages, all details of the procedure were the same as in Study 1a. After the tact review trials, the test screen was placed on the table, to reduce possible cueing. Experimenter 1 conducted the test with familiar stimuli. The child's listener behaviour was tested by spinning the carousel and asking, "Which one is the Reptile (or Fish/Fruit/Vegetable)?" No reinforcement was given during the test.

Arbitrary stimulus tact training and listener behaviour test. Following pre-training and testing with the familiar stimuli, the complete procedure of Study 1a was implemented. Again, for each individual child the eight alien animals were randomly assigned to four two-member categories, denoted by the nonsense names "hib", "feb", "tor", and "lup", and divided over two sets, each consisting of one hib, one feb, one tor and one lup. The relations between the specific animals and the trained names for each participant in Study 2a are shown in Table 3.2.

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 100%. Similarly, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Table 3.2 For each child, the names allocated to each alien.



Results

Tact training—Familiar stimuli: Set 1, Set 2, mixed sets, and reinforcement reduction. All 13 children completed pre-training with the familiar stimuli, for which the minimum number of trials was 216. The children needed on average 295 trials (range 216—636). For the 4 females the average was 243 trials (range 216—288), for the 9 males it was 319 (range 216—636). Without the extreme number of 636 trials, the average for the males was still 279 trials. Figure 3.2 shows the total number of trials for the individual children. As can be seen, 3 children (1 girl, and 2 boys) only needed the minimum of 216 trials. Ten of the 13 participants had between 200 and 300 trials, while 3 boys had more than 300. Adam required 636 trials, considerably more than the other children. For the children who needed (many) more trials than the minimum, most errors were due to their failure to discriminate between the fruits and vegetables. This was also the case for Adam. Because at some point he did not seem to make any progress, he was given additional sessions, as described below.

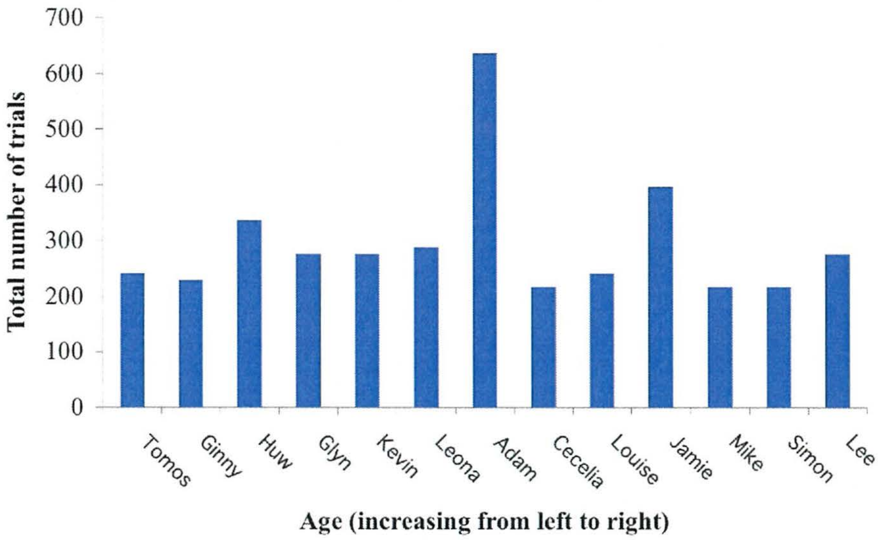


Figure 3.2 Total number of trials per child in tact pre-training (all stages combined).

Listener behaviour test—Familiar stimuli. After reaching criterion for tact training with the familiar stimuli, all children went on to pass the test for listener behaviour.

Additional training sessions—Familiar stimuli. Adam was the only child who was given additional training sessions with the familiar stimuli. Although he reached criterion for Set 1 within 60 trials, and for Set 2 in 84 trials, during mixed sets training his tact responses to the fruit and vegetable stimuli deteriorated. He had 228 trials (132 trials more than the minimum) for that stage, and then he moved on to the mixed sets no reinforcement stage. When he did not reach criterion in that stage, additional sessions with modified procedures were conducted with him.

In these sessions, only the fruit and vegetable stimuli were targeted, and he was told at the start of the session that fruits are for pudding and vegetables are for dinner. After each tact response, he was asked why it was a fruit/vegetable, with the added question, "Is it for dinner or pudding?" After two 12-trial blocks, the procedure was changed, so that on each trial, Adam was asked, "What's this?"

(Adam would provide the lower-level name), followed by the question, “And is it for dinner or pudding?” After Adam’s response to this he was asked, “So it’s a ... ?” (Adam would provide the appropriate higher-level name). In four 12-trial blocks his tact performance improved markedly (to 11/12 correct, but then regressed to 9/12).

Then, in addition to this new procedure, a new type of reinforcement was introduced: for every correct response, the experimenter held up one finger. And once the child had given three correct responses (Experimenter 1 holding up three fingers), he could press a button on the Thomas the tank engine book (a music book — pressing a button on this book produced a sound). The number of correct responses increased, and then remained stable: 11/12, 9/12, 10/12, 12/12, 12/12, 11/12. This was done over six 12-trial blocks, but after the first two blocks, the prompt was changed back to the usual, “What’s this?” leaving out the added questions.

Interestingly, apart from Adam’s improved performance, another change was observed. Whereas before introduction of the music book reinforcer he was asking after almost every trial, “Are we finished?” and he could not wait to read a story, now he could not wait to start playing the game, and he wanted to play more of it, instead of reading the story! He was not just enjoying the pressing of the buttons on the music book, but also watching the fingers going up. He showed excitement when two fingers were held up, and he only needed one more to press a button. Furthermore, before introduction of the music book, he moved back and forth on his chair during sessions and continually looked around the room, whereas after introduction of the book he sat still, all through the sessions, completely focused on the stimuli on the carousel and on the experimenter's fingers that she held up.

Then, still with use of the music book reinforcement, four 12-trial blocks of normal training were conducted, in which all stimuli were targeted, and not just the fruits and vegetables; the number of correct responses in each block were: 12/12, 12/12, 12/12, 11/12. Eventually, four 12-trial blocks for the mixed sets no reinforcement stage were conducted (12/12, 12/12, 12/12, 12/12). In this stage, no fingers were held up for correct responses. Adam did not like this, he said: “I want you to count on your fingers”. Experimenter 1 responded: “I can’t, because Peter Penguin (hand puppet) is on my hand, but I’m counting in my head and you get to press the buttons afterwards”. After each sheet, Adam was permitted to press a few buttons, but this was not contingent on particular responses. He was concentrating very well and clearly thinking carefully about the fruits and the vegetables, starting his response in these particular trials with “vvv…./fff…”, then thinking a bit longer, before giving his full answer. In the listener behaviour test for the familiar stimuli Adam made no errors in 48 trials. Adam had received 31 12-trial blocks for Set 1, Set 2 and the mixed sets with reinforcement, taken together. In addition, he had 22 12-trial blocks for the mixed sets under no reinforcement, including the additional training sessions. This made a total of 636 trials.

Tact training—Arbitrary stimuli: Set 1, Set 2, mixed sets, and reinforcement reduction. Twelve of the 13 participants completed the alien tact training. Only Glyn did not. After 432 trials of the alien tact training, the sessions with him were discontinued because he did not want to play any more.[1] The children who did complete alien tact training needed on average 397 trials (range 252—600). As for a

---

[1] In pre-training, Glyn was slightly below average with regard to the total number of trials, but he learned to tact the aliens in Set 1 in just 84 trials – of all participants, only Mike took less trials. Glyn then needed three times the average number of trials to learn Set 2, and finally, because he kept making errors on the 96 trials of mixed sets with reinforcement, especially with the hibs and the lups, he did not want play the game again. The sessions were discontinued at this point.

breakdown of the number of trials for the separate stages, for Set 1 the average was 147 (range 60—228), while for Set 2 it was 86 (range 36—252). For the mixed sets with reinforcement it was 104 (range 96—132), and for the mixed sets under no reinforcement it was 65 (range 48—192). The total number of trials to complete training for the 4 females was on average 510 (range 384—600), while for the 9 males the average was 347 (range 252—432).

The data discussed so far were group averages. The total number of trials for each participant, including the child who did not complete training, is shown in Figure 3.3. The number of trials is presented as a function of age, and there is no indication that the older children needed fewer trials to criterion than the younger ones. It can be seen in Figure 3.3 that Mike's tact performance met the criterion in 252 trials, whereas Ginny needed 600 trials.

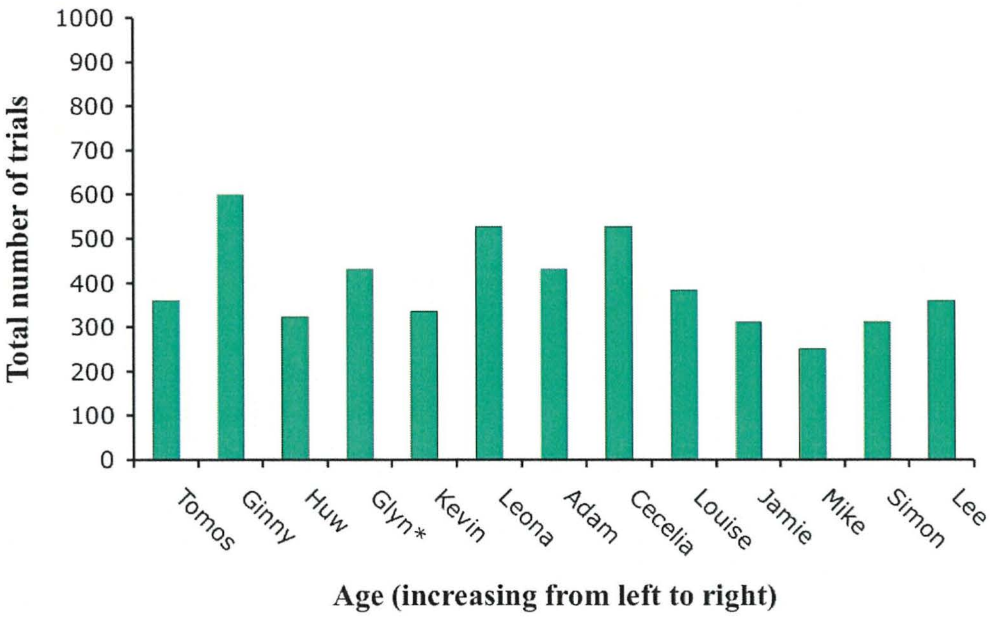


Figure 3.3 The total number of trials needed to reach criterion over all stages in alien tact training until the listener behaviour test, for all children, with the youngest child on the left (Tomos, 3/3) and the oldest child on the right (Lee, 4/4). For all ages: see Table 3.1. (*Child did not complete the study.)

Figure 3.4 shows the number of trials for female and male participants, respectively. As a group, the males needed fewer trials to criterion (range: 252—432) than the females (range: 384—600). Only for 1 girl, Louise, was the number of trials (384) within the range for the boys. However, there were twice as many boys as girls so this outcome may be due to chance.

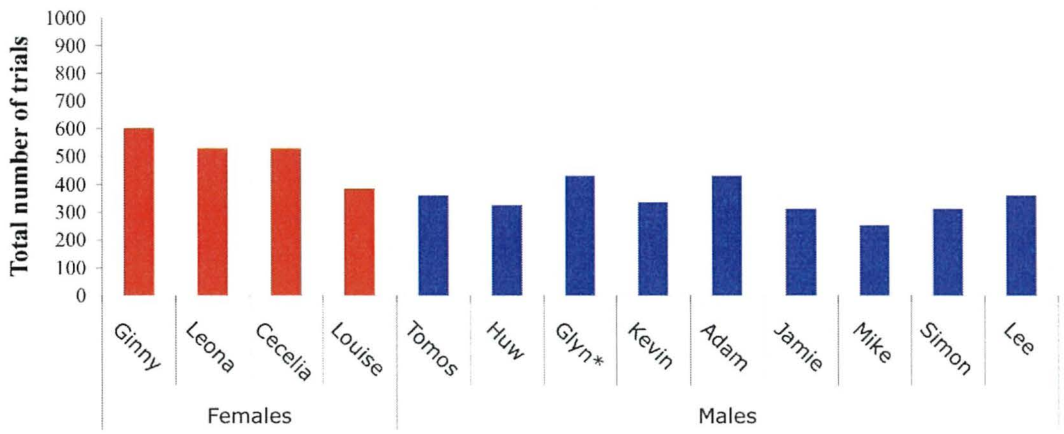


Figure 3.4 Total number of tact training trials for all females and all males.

Figure 3.5 shows for each child the number of trials to criterion for each stage within the study. Note that pre-training with familiar stimuli involved the same stages as tact training with the alien animals, although the number of trials for each stage is not specified separately. For alien tact training, the separate stages are specified. It can be seen that in alien tact training for Set 1 Mike only needed 60 trials (24 more than the minimum), while Glyn had just 84, and Simon and Lee had 96 trials each. Six children had between 100 and 200 trials, while the remaining 3 children had between 200 and 230 trials. For Set 2, most children (10 out of 13) required (far) fewer trials than for Set 1. When presented with Set 2, the child was already used to being presented with newly designed alien animals, and to the names (both sets consisted of animals named hib, feb, tor, and lup). Four children (1 girl, 3 boys) reached criterion for Set 2 in the minimum of 36 trials, while another 4 boys had just one 12-trial block more. Three children required less than 100 trials, 1 child needed 120 trials, and 2 other children had well over 200 trials. At the stage of mixed sets with reinforcement, 8 children required no more than the minimum of 96 trials, while 2 other children needed just one more 12-trial block. In the mixed sets no reinforcement phase, 10 children met the criterion in the minimum number of trials (48).

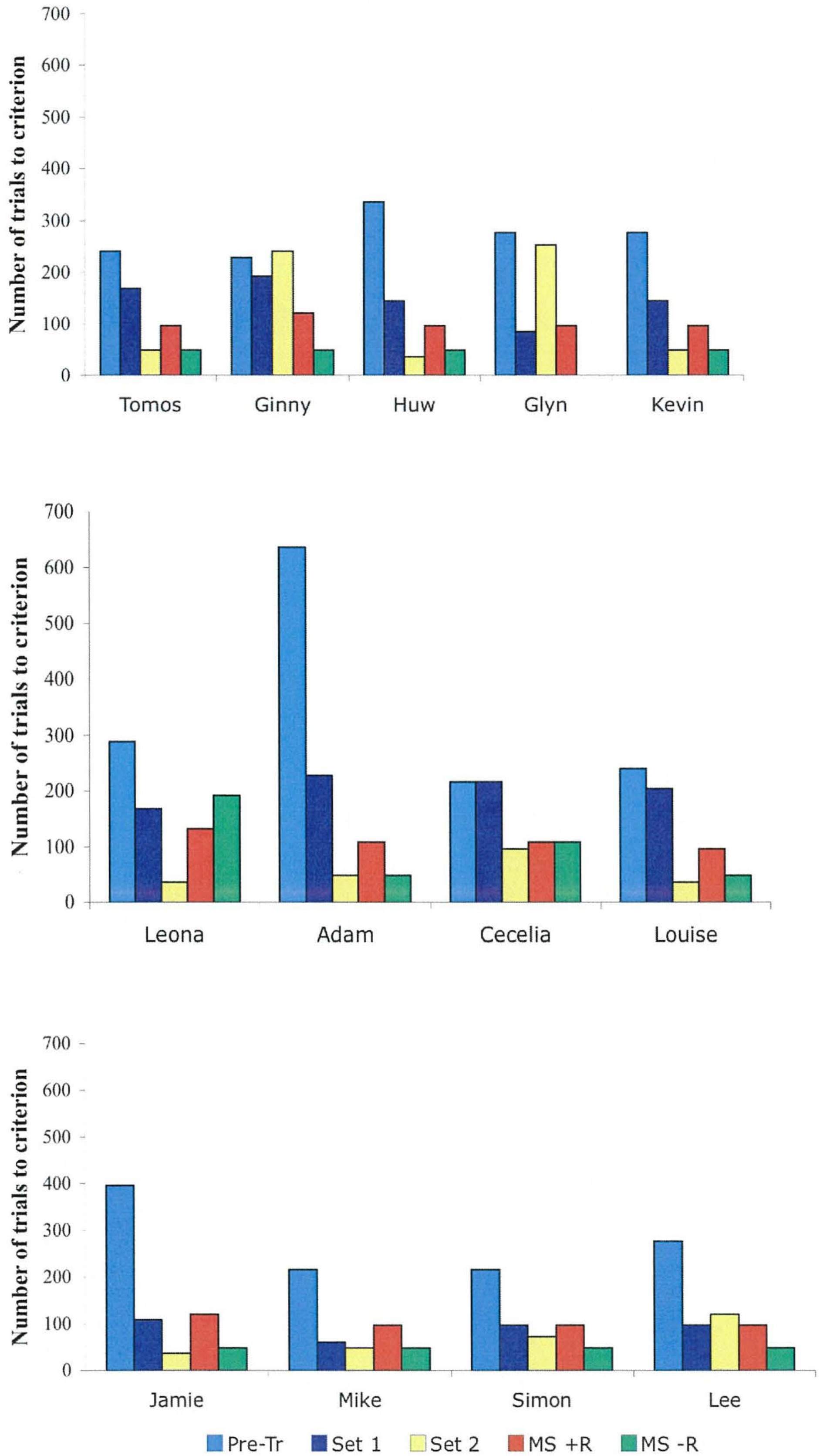


Figure 3.5 For each child, number of tact training trials to criterion in Pre-training with the familiar stimuli, and for Set 1, Set 2, mixed sets with reinforcement (MS+R), and mixed sets without reinforcement (MS-R) with the alien stimuli.

Listener behaviour test. Figure 3.6 shows the percentage of correct responses per child for each of the two mixed sets and for the two test sheets (a and b) per set. Pass level was 58% (14/24 trials) correct for each set. All 12 participants passed the test; 6 children performed without error over all 48 trials (24 trials per mixed set), 2 children made one error, and 3 children made two errors. The remaining child, Cecelia, made seven errors mostly on trials for Feb 2, a stimulus for which she also showed errors in tact training.

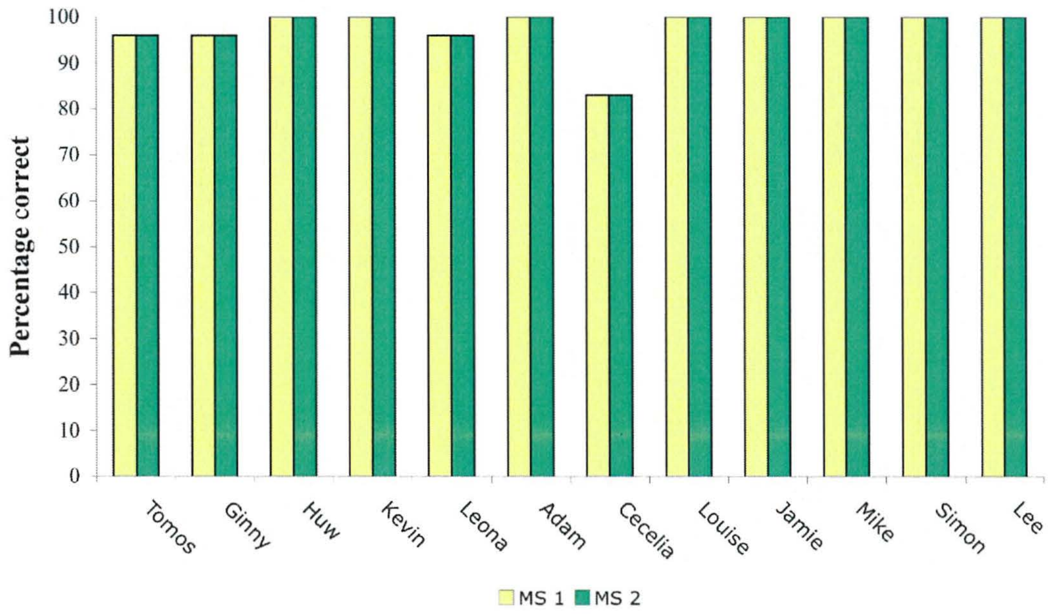


Figure 3.6 For each child, the percentage of correct responses in the listener behaviour test for Mixed Set 1 (yellow bars) and Mixed Set 2 (green bars). Pass level was 58% (14/24 trials) correct for each set.

Table 3.3 lists the vocalisations produced by the children during the listener behaviour test at the lower name level. Appendix C contains a collection of comments that the children made during training sessions of all studies.

Table 3.3 The children's vocalisations during the listener behaviour test at the lower name level.

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Tomos	Mixed Set 1	Which one is the hib?	(points at hib) He's got one eye.
	Mixed Set 2	Which one is the hib?	(points at lup) Lup, (points at hib) hib
Kevin	Mixed Set 2	Which one is the tor?	(points at tor) It's a funny one!
Jamie	Mixed Set 2	Which one is the hib?	(points at hib) He's a kangaroo!
		Which one is the feb?	(points at feb) She's a naughty one too.
		Which one is the tor?	(points at tor) He's biting my finger off.
		Which one is the lup?	(points at lup) He's nice.
		Which one is the hib?	(points at hib) Lup (E: Which one? Jamie points at hib again)
Simon	Mixed Set 2	Which one is the hib?	(points at hib) It has a horn.
Lee	Mixed Set 1	Which one is the feb?	Feb (then points) x3
		Which one is the hib?	Hib (then points) x2
		Which one is the tor?	Tor (then points)
	Mixed Set 2	Which one is the hib?	Hib (then points) x2
		Which one is the lup?	Lup (then points) x2
		Which one is the feb?	Feb (then points) x2

Discussion

After Study 1a, Study 2a found further confirmation that children could learn to tact (i.e., see animal – say "name") the aliens that were newly designed for these series of studies. The data of Study 2a show that all 13 children who participated reached criterion for tacting the aliens in both Set 1 and 2, while 12 of them reached criterion for all stages in the mixed sets tact training procedure. As concluded in Chapter 3, this suggests that the alien animals are suitable experimental stimuli, also for further studies.

Furthermore, Study 2a, like 1a, focussed on learning of name relations, and on the relation between tact behaviour and listener behaviour. It aimed to test the naming account in the same way as Study 1a, by training of tact behaviour for the eight alien animals, and then testing for listener behaviour. Would children who had learned to tact the alien animals, demonstrate the corresponding listener behaviour on hearing the object name (i.e., hear /name/ – select animal), without being directly trained to do so? The data show that 12 of the 13 participants reached criterion for all stages of tact training and all passed the listener behaviour test.

This extends the support that was found for the naming account in Study 1a. As was noted there, the naming account of Horne and Lowe (1996) suggests that when a child is trained to tact an object, in the course of this the child is likely to also learn the corresponding listener behaviour, without it being separately trained. Cecelia, made more errors on the test than the other participants; the errors concerned a stimulus with which she also showed errors in tact training, which suggests that both her speaker and her listener behaviour with regard to this animal were unstable. All in all, 12 children in Study 2a learned to name the aliens (i.e., they learned to tact them, and show appropriate listener behaviour towards them). In several of the test trials, Lee echoed the listener stimulus out loud before producing the listener response, which was also observed in some of the children in Study 1a.

A final aim of Study 2a was to determine whether the children would require less tact training trials for the alien stimuli than the children in Study 1a, given that in Study 2a the children were first exposed to tact training with familiar stimuli whereas participants in Study 1a had no such pretraining (and see Horne, et al., 2006, p. 268). A first comparison of the data of studies 1a and 2a reveals that the average number of *alien* training trials needed by children who had had pre-training, was about 400, which was much lower than the average of about 600 trials for children who had not had pre-training. A detailed comparison of both studies can be found in the general discussion (Chapter 4). In Chapter 2, it was discussed that presenting four stimuli at a time made far fewer trials necessary for the children to learn to discriminate between the aliens, than in pair wise training. It can now be concluded that the new method of presenting four stimuli at a time, in combination with pre-training, gave the best result; that is, the fewest number of trials were needed for alien training in that case. However, when considering the total trials spent in experimental sessions, the number of trials saved in alien training is more or less the same as were spent on pre-training: the children who had pre-training may have had on average 200 trials less in alien training than the children who did not have pre-training, but the minimum number of trials in pre-training was 216. When considering the children's learning to tact the aliens, it is, of course, interesting to find that prior experience of the procedures with familiar stimuli facilitates the children's tact performances with the alien stimuli, even though both the procedure and instructions may seem simple and straightforward.

Study 2b – Learning higher-level names

Study 2b, like 1b, investigated one way that children may learn higher-level names for objects. Is it perhaps enough for them to learn the intraverbal relation between the lower-level and higher-level names, or do all these names have to be trained directly in the presence of each of the objects?

To investigate the first possibility, in Study 1b a partly echoic and partly intraverbal word game was introduced, linking each of the lower level alien names (hib, feb, tor, lup) with one of two higher-level names (zaag, noom).[2] Just as dogs and cats are called animals at a higher level of abstraction, the names "zaag" and "noom" were chosen as potential higher-level names for the aliens. Study 2b was the same as 1b, apart from the fact that pre-training was added. That is, before the word game was played with the alien names, the same game was played with the *familiar names*, following exactly the same procedure as the alien name word game. Once the children had learned the intraverbal name relations, which linked the lower-level familiar names (reptile, fish, fruit, vegetable) with the higher-level familiar names (animal, food), they were tested for listener behaviour with regard to these higher-level familiar names (i.e., hear /higher-level name/ – select shape), before moving on to the alien word game.

The aim of Study 2b was to see whether more children would show correct listener behaviour at the higher name level with regard to the aliens, when pre-training with the names of the familiar stimuli was added, providing a real life parallel and ensuring that the child would respond appropriately to the procedure and the instructions.

---

[2] Different name links were randomly assigned to individual children.

Method

Participants

Of the participants in Study 2a who passed the test for listener behaviour, 8 children (2 females and 6 males) were available for part 1 of Study 2b. They were between the ages of 3 years and 4 months, and 4 years and 9 months at the start of the study (see Table 3.4). For the second part of this study, Lee was not available, because he had left the nursery.

Table 3.4 Participants' sex, age at start of pre-training, along with age at start of alien word game training, and age at listener behaviour test, for Leg 1 and 2 with the aliens.

Participant	Sex	Age at start pre-training*	Age at start alien training Leg 1	Age at testing Leg 1	Age at start alien training Leg 2	Age at testing Leg 2
Tomos	M	3/4	3/6	3/7	3/7	3/10
Huw	M	3/5	3/6	3/6	3/6	3/6
Ginny	F	3/8	3/9	3/11	3/11	4/0
Adam	M	3/9	3/10	3/10	3/10	4/0
Louise	F	3/10	3/11	4/0	4/0	4/2
Mike	M	3/11	3/11	4/0	4/0	4/0
Simon	M	4/5	4/5	4/6	4/6	4/6
Lee	M	4/9	4/10	4/10		

* Ages in years/months; age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).

Apparatus and Stimuli

The experimental setting in both parts of Study 2b (the word game with familiar names, and with alien names) was as in Study 1a, except that there were no carousels and no alien animal stimuli on the table during word game training trials. Separate stimulus sheets for pre-training and for the alien name word game listed the verbal stimuli to be provided by the experimenter, and were used to record the child's behaviour during the session. In the listener behaviour test sessions at the end of pre-training the eight familiar stimuli described in Study 2a were used, while in testing at the end of the study the alien animal stimuli were used. The listener behaviour test sessions with the aliens were conducted by Experimenter 2, who used stimulus sheets that specified the position of the aliens on the table and the order of targeting without linking the higher-level names and the stimuli, so that she remained blind to the trained intraverbal relations and could not cue the child's responses during the test trials. In all listener behaviour test sessions, with familiar shapes and alien animals, the wooden test screen was used (see Study 1a). During training sessions the main reinforcers were social praise and a story at the end of each session. In addition, sometimes a music book and magic cards were used (see Study 2a).

Procedure

This was a single case design with replication across participants. No parents were present during the sessions. There were one or two sessions per day with the child, taking about 10 to 20 min each. For each participant, Study 2b was started as soon as possible after completion of Study 2a.

Echoic training with familiar names. In pre-training, a word game as described in Study 1b linked the names of the familiar stimuli at the lower level (reptile, fish, fruit, vegetable) to their names at the higher-level (animal, food). Counterbalancing took place with regard to the order in which the links were trained for individual children. So one child could start, for example, with the reptile-animal link, followed by the fruit-food link, while another child could start with the vegetable-food link, followed by the fish-animal link. Apart from the names that were used, the procedure was the same as in Study 1b.

First, the echoic game, called *You say what I say*, was played, in which the child echoed an increasingly complex verbal stimulus produced by the experimenter. If in the first of a block of eight trials the verbal stimulus was "reptile", then it would be "reptile-animal" in the second trial, followed by "reptile-animal-reptile-animal" in the remaining six trials. In all eight trials in the second block, the verbal stimulus was also "reptile-animal-reptile-animal". Stimulus sheets listed the verbal stimuli to be provided by the experimenter.

The child's correct echoic responses were followed by social praise (e.g. "Yes!", or "Well done"), while for incorrect echoic responses, the experimenter provided corrective feedback by producing the verbal stimulus again.

Intraverbal training. The second part of the verbal training was an intraverbal game, in which the child responded with a "chain" of lower-level and higher-level names upon hearing just the lower-level name. The child was instructed that whenever the experimenter said, for example, "reptile", the child should say "reptile-animal-reptile-animal". The experimenter provided the same verbal stimulus over a minimum of two blocks of eight trials each. Praise or correction followed the child's responses.

Reduction in reinforcement rate. Once criterion was reached for the second part of the verbal training, a further two blocks of eight trials were conducted, with verbal stimuli identical to the stimuli in the intraverbal game. This time the trials were presented under zero reinforcement, to prepare the child for the zero reinforcement conditions during the test sessions, and thereby to ensure that responding would continue also under extinction. In these no reinforcement sessions the experimenter reintroduced the penguin hand puppet (Peter Penguin), and asked the child to teach Peter the word game. At this stage, as before, the learning criterion was seven out of eight correct within a block over two consecutive blocks. If the performance of any child did not meet the zero reinforcement criterion, then reinforcement was reintroduced until responding was once again at criterion. The child was then tested once again in extinction, and so on, until the child's performance met the zero reinforcement criterion.

All three stages in the word game were played with respect to both higher-level categories. During the word game training period, maintenance training of the previously trained tact relations (familiar stimuli in Set 1 and 2) was conducted. There were no familiar stimuli on the table during the word game.

Listener behaviour test for higher-level familiar names. This test prompted the child to show listener behaviour with regard to the higher-level names. Because children of this age, already know that crocodiles, snakes, sharks and stingrays are animals, and that apples, bananas, mushrooms and carrots are foods, the listener behaviour test for the familiar stimuli did not determine whether the echoic and intraverbal training had resulted in the child being able to make listener behaviour discriminations (hear /name/ – select shape) with regard to the higher-level names. Instead, it just served to provide the children with an introduction to the listener behaviour test with the aliens that would follow later, that is, after the alien word game. In the test sessions the experimenter started with review trials for lower-level tacting of the familiar stimuli. Following that, the intraverbal links were reviewed. During review trials, occasional errors were corrected. After the review trials and before the test, the experimenter instructed the child as in Study 1b: "Now that you've learned the names of the animals, and the word game, we'll see how that can help you on the next game."

The test after pre-training was conducted by Experimenter 1. The test screen (see Study 1a) was placed on the table between the experimenter and the child. Testing was done by putting two of the familiar stimuli (each of a different higher-level category) in front of the child and asking, "Where's the animal/food?" For the test, four familiar stimuli were used, belonging to two lower-level categories and to two higher-level categories (for example, the two reptiles and two fruits could be used for the first test for one particular child). The number of trials and blocks of trials, the way of presenting and targeting, and the listener behaviour criterion were exactly as in Study 1b. No reinforcement was given during the test, nor was behaviour corrected during trials. Clues, as given in Study 1b, were not necessary with the familiar stimuli.

After testing for listener behaviour at the higher name level with regard to the first two links (e.g., reptile-animal and fruit-food), the intraverbal training procedure was conducted for the remaining links (e.g., vegetable-food and fish-animal), followed by another test for listener behaviour at the higher name level.

Echoic training with arbitrary stimulus names. Following pre-training and testing with all eight familiar stimuli, the complete procedure of Study 1b was implemented with the arbitrary (alien) stimulus names.

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 100%. In the same way, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Familiar stimuli: echoic training, intraverbal training, and reinforcement reduction. All 8 participants completed the echoic and intraverbal training with the names of the familiar stimuli for the first leg of the hierarchy (that is, two out of four links — see Study 1b, for an explanation of the legs of the hierarchy). Figure 3.7 shows the number of training trials for Leg 1 and Leg 2 in pre-training. The children needed on average 121 trials (range 96—160) to get through Leg 1 of pre-training; 35 trials for the echoic training (range 32—56), 53 for the intraverbal training with reinforcement (range 32—80), and 33 for intraverbal training without reinforcement (32—40 trials). The females had 156 trials on average (Ginny had 152 trials, Louise had 160), while the average for males was 109 (range 96—144). Unlike the alien word game of Study 1b, 3 children needed considerably more intraverbal training, than the minimum of 32 trials. This concerned 1 boy (Tomos, with 80 trials), and the 2 girls (with 72 and 80 trials, respectively). After intraverbal training for Leg 1 of the hierarchy, the children were tested for listener behaviour at the higher name level for Leg 1, with the familiar stimuli.

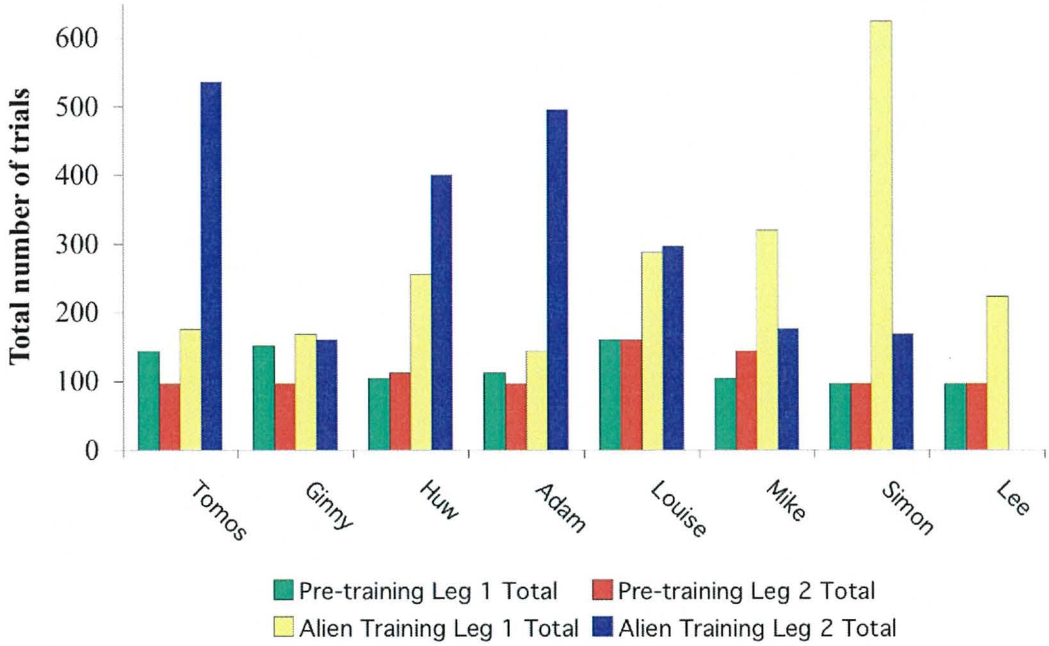


Figure 3.7 The total number of training trials for all individual children in pre-training, that is, the word game with the familiar names (Leg 1 in green, Leg 2 in red), and in the word game with the alien names (Leg 1 in yellow, Leg 2 in blue).

Word game pre-training, Leg 2: echoic training, intraverbal training, and reinforcement reduction. All 8 children then moved on to, and completed, word game training with the names of the familiar stimuli for Leg 2 of the hierarchy. The average number of trials for Leg 2 was slightly lower than for Leg 1 of pre-training: 112 trials were needed (range 96—160); 32 trials for the echoic training (all children required the minimum of 32 trials), 48 for the intraverbal training with reinforcement (range 32—96), and 32 for intraverbal training without reinforcement (all children only required the minimum of 32 trials). Females had on average 128 trials (Ginny had 96 trials, Louise 160), while males had 107 (range 96—144).

Listener behaviour test at the higher name level with familiar stimuli. Following training for the first two links of familiar names, the children were tested for Leg 1 of the hierarchy. After training for the remaining two name links, the children were tested for Leg 2. All 8 participants passed the test for listener behaviour at the higher name level with the familiar stimuli, for both Leg 1 and Leg 2.

The word game with the alien names, Leg 1: echoic training, intraverbal training, and reinforcement reduction. With all 8 participants the word game was then played with the alien names. All children completed the echoic and intraverbal training for the first leg of the hierarchy. On average they needed a total of 275 trials (range 144—624); 52 trials for the echoic training (range 32—120), 190 for the intraverbal training with reinforcement (range 72—472), and 33 for intraverbal training without reinforcement (range 32—40). The total number of training trials for females was on average 228 (Ginny had 168 trials, Louise 288), and for males it was 291 (range 144—624). So males needed more trials than females. Three out of 8 participants (all males) needed no more than the minimum number of trials to reach criterion in the echoic training stage, while in intraverbal training without reinforcement this was the case for 7 participants (1 boy needed one more 8-trial block). In intraverbal training with reinforcement all children needed many more trials than the minimum of 32. After completion of training for Leg 1 of the hierarchy, the children were tested for listener behaviour at the higher name level for Leg 1, with the aliens.

The word game with the alien names, Leg 2: echoic training, intraverbal training, and reinforcement reduction. Seven children started the word game training with the alien names for Leg 2 of the hierarchy, after being tested for listener behaviour at the higher name level for Leg 1 (1 child, Lee, left the nursery at the end of the Leg 1 procedures). All seven completed the training. On average they needed 319 trials in total (range 160—536); 61 trials for the echoic training (range 32—112), 221 for the intraverbal training with reinforcement (range 96—392), and 38 for intraverbal training without reinforcement (range 32—56). Therefore, on average the children needed 44 trials more than for Leg 1, but at the individual level the differences between the number of trials for both legs were much bigger, at least for some children. Tomos, for example, went from 176 trials in Leg 1 to 536 for Leg 2. Huw went from 256 to 400 trials, and Adam from 144 to 496. However, Mike and Simon showed a reverse pattern. Mike had 320 and Simon 624 trials for Leg 1, while for Leg 2 they had 176 and 168 trials, respectively. The females had exactly the same average for Leg 2 as they had for Leg 1, 228 trials (Ginny had 160 trials, Louise 296). Males needed 355 (range 168—536). So this time females needed considerably less trials than males. For Leg 1 they also needed fewer trials than the males, but the difference was 63 trials, while for Leg 2 females needed 127 trials less than the males. Three participants needed no more than the minimum of 32 trials in echoic training, while for 5 participants this was the case in intraverbal training without reinforcement. In intraverbal training with reinforcement all children needed (many) more than the minimum of 32 trials.

Pre-training and the alien word game: the number of training trials compared. Apart from showing the number of training trials for Leg 1 and 2 with the familiar stimuli in pre-training, Figure 3.7 also showed the number of trials for Leg 1 and 2 in training with the alien names, per child. When comparing the number of trials needed in training for both legs of the alien name word game, with the number of trials needed in pre-training, a difference is found. For pre-training, all children needed between 96 and 160 trials for each of the two legs. In training for the alien name word game, however, only Ginny needed a similar number of trials for each of the two legs; all other children required considerably more trials for at least one of the two legs. For Leg 1 of the alien name word game, 3 children have a number of training trials that is comparable to the number of trials in pre-training (Tomos, Ginny, and Adam), while the others needed at least double the number of trials they had in pre-training for that leg, and for 1 child (Simon), six times that number of trials were required. In Leg 2 of the alien name word game, numbers of training trials were comparable to those required in pretraining for 3 children (Ginny, Mike, and Simon); in the case of the remaining children, Louise needed about double the number of trials required in pre-training for that leg, while Huw required four times as many, and Tomos and Adam, five times as many trials.

Listener behaviour test at the higher name level. Figure 3.8 shows the percentage of correct responses per child for each of the zaags and nooms in the test for listener behaviour at the higher name level, for Leg 1 and Leg 2 of the hierarchy. The listener behaviour criterion was 10 out of 12 (83%) correct per alien. Following training for the first two name links, a child was tested for Leg 1 of the hierarchy.

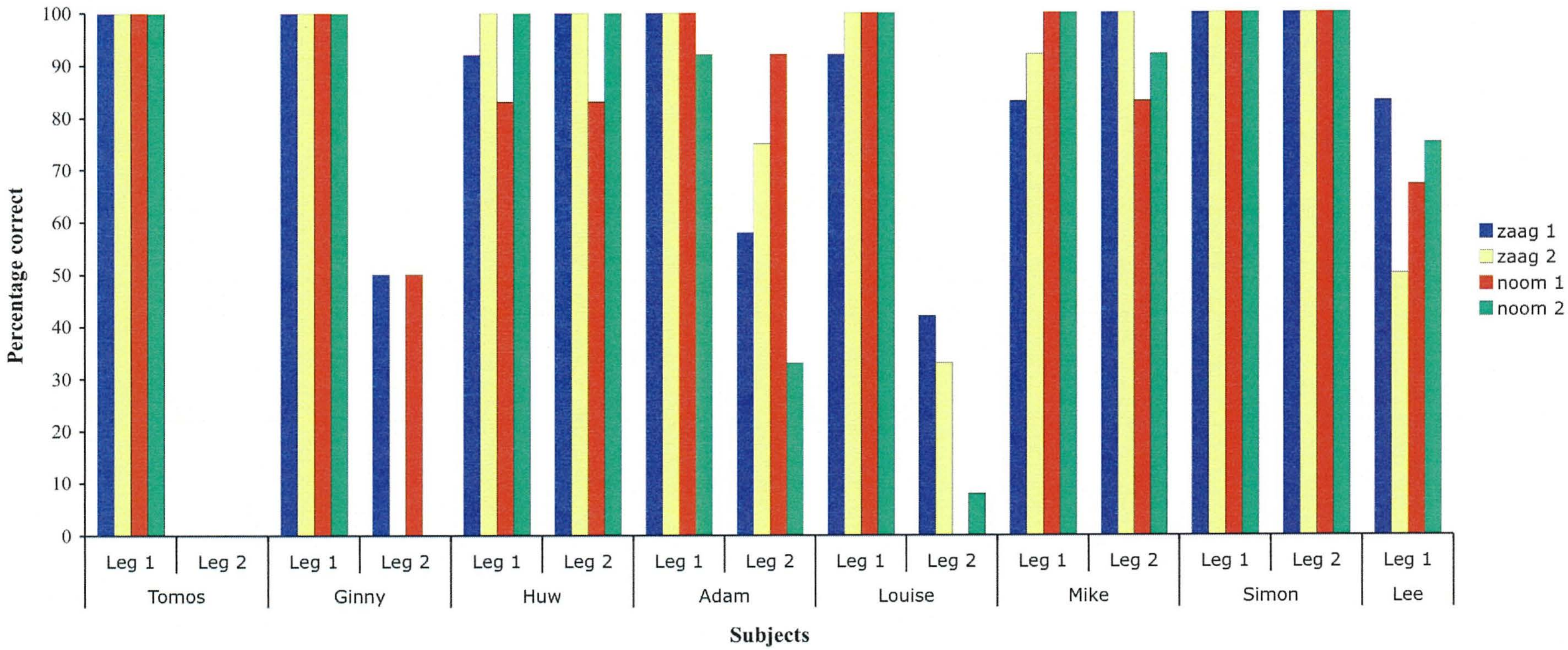


Figure 3.8 The percentage of correct responses per child for all the zaags (in blue and yellow), and nooms (in red and green), separately, in the test for listener behaviour at the higher name level. The listener behaviour criterion was 10 out of 12 (83%) correct per alien.

Leg 1. As can be seen in Figure 3.8, of the 8 participants given the listener behaviour test for Leg 1 of the hierarchy, 7 passed. Three of them (Tomos, Ginny, and Simon) performed without error on the 48 test trials (12 trials per animal). Two children (Adam, and Louise) make one error, while 2 other children (Huw, and Mike) make three errors each. One child (Lee) failed the test. He reached the criterion of 83 percent correct for only one of the four animals. For the other three animals he scored between 50 and 75 percent. After the test, Lee was given follow-up sessions and retests, described after the results for Leg 1, and eventually he passed. The vocalisations produced by the children during testing for Leg 1 are listed in Table 3.5.

Table 3.5 The children's vocalisations during the listener behaviour test at the higher name level, Leg 1.

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Adam	Hib 1 + Tor 1	Which one is the zaag?	Tor!
	Hib 1 + Tor 2	Which one is the noom?	Both. (E: Only one, which one?)
	Hib 2 + Tor 1	Which one is the noom?	(picks up correct animal) hib
Louise	Feb 1 + Hib 1	Which one is the noom?	Don't know. (E: Just try ... do what you think.)
		Which one is the noom?	Don't know. (E: But you were trying before. Try again.)
Simon	Hib 2 + Lup 2	Which one is the noom?	(takes a while to pick, then explains:) These ones have a different name.
		Which one is the zaag?	The zaag is a dog! Lup (then points correctly)
		Which one is the noom?	(points correctly) Yes, that one because hib-noom.

As in Study 1b, after testing, one or more trials were usually repeated, and following a response the child was asked why that response was given. So the child was asked: "Why is that a zaag/noom?" As in Study 1b, some children replied that they didn't know, or said, “ 'Cause it is the noom” (Lee). Tomos came up with a variation of this. He said: “Because it’s like a zaag …”. Then he explained: “It’s got spikes”. But not all Tomos’ zaags had spikes. Other children also referred to physical features of the animals. Louise said: “It has bumps on its head”. And Adam referred to supposed behaviour of the animal: “ 'Cause it pokes”. For another animal, Adam said: “ 'Cause it’s got a wiggly”. Only Simon replied by referring to the lower-level name: “ 'Cause it’s a lup/ hib”.

Leg 2. After the test for Leg 1, the children were trained for the other two word links (Leg 2). When training was completed, the children were tested for Leg 2. Of the 7 participants who were tested, 3 passed the test, whereas all 7 of them had passed the test for Leg 1. One child, Simon performed errorlessly on the test for Leg 2 (he also produced no errors on the test for Leg 1). Huw made two errors, while Mike made three. Of the children who failed the test, Adam did reach criterion for one animal, and had only one error too many for another animal. However, for the other two animals he scored 7/12 and 4/12, well below the criterion of 10/12. Louise showed reversal of responses for two animals, and scored at chance level, or below, for the other two animals (42 and 33 percent correct, respectively, per animal). Louise’s scores per sheet give more information. Her scores were 0/12, 1/12, and 0/12, so showing almost complete reversal of responses over the first three sheets, and then she was given the final and most explicit clue (i.e., A hib *is* a zaag), and she went on to score 9/12 for the last sheet. Louise had follow-up sessions (as described at the end of this results section), and was retested. Very interesting results are those for Ginny and, even more so, for Tomos. Ginny showed reversal of responses for two animals, and had 50 percent correct for the other two. She had all 12 trials correct for sheet 1, and then gave the incorrect response for the 36 remaining trials, despite the clues that were given at the end of the trial blocks. Tomos showed complete reversal of responses for all four animals. His responses were incorrect in all 48 trials. Both Ginny and Tomos had follow-up sessions (see below), and were retested. During testing for Leg 2 none of the children produced any vocalisations.

In the extra trials with added why-questions, Ginny first replied, “ 'Cause it is”. When prompted further (“But can you think why?”), she said: " 'Cause it’s got spikes”. But the two animals presented at that time were Tor 2 and Hib 1, the first of which has spikes on top, and the other alien has spikes on the side.

Additional sessions. Children who failed the test for Leg 1 or Leg 2, were given additional sessions if still available for testing. This concerned Tomos, Ginny, and Louise. Sessions with Adam were discontinued at this point, due to illness.

Tomos failed the test for Leg 2, even though he was one of 3 children whose performance was errorless in testing for Leg 1. In the test for Leg 2 he showed complete reversal of responses (scoring 0 for all aliens). He then had additional sessions, starting with a session of maintenance training for all components, with the familiar stimuli (tacting and word game), and with the alien animals (tacting and word game). Tacting and word games were still in place. Following this, Tomos was retested. His scores were exactly the same as before (0 for all aliens). It should be stressed that during testing Tomos appeared to attend well to the task and responded promptly on each trial. However, he consistently picked the wrong alien (on one trial he gave a correct response, but just as the aliens were taken off the table, he "corrected" himself). After this retest, Tomos had four extra training sessions in which it was clear that the trained tacting and intraverbals were still in place. Then the word game "You say what I say" was played, with an extension. Rather than just providing the models "lup-zaag-lup-zaag" and "hib-noom-hib-noom", to be echoed by the child, the experimenter said: "lup-zaag-lup-zaag ... a lup is a zaag" and "hib-noom-hib-noom ... a hib is a noom" which the child echoed. Four 8-trial blocks (32 trials) were conducted per link, making a total of 64 trials. Finally, he was retested one more time. Again, he showed reversal of responses (he gave one correct response for Noom 1, but otherwise all responses were incorrect).

Ginny also failed the test for Leg 2. Like Tomos, she was one of the 3 children who performed without error in the test for Leg 1. In the test for Leg 2, she showed reversal of responses for two animals, and she had 50 percent correct for the other two. After the test for Leg 2, training for the word game was resumed (the normal intraverbal game: "What do you say when I say hib/lup/tor/feb?"), combined with maintenance training for tacting the aliens. Ginny started to make errors in the hib-zaag link (saying "hib-noom"). Because of this, training went on for 14 sessions, in which she regularly made the same mistake. When retested, she showed complete reversal (0/48 trials correct). She then received additional training: the new "You say what I say" game with extension was played. Experimenter 1 said, "hib-zaag-hib-zaag ... a hib is a zaag" and "tor-noom-tor-noom ... a tor is a noom" which the child echoed. Six 8-trial blocks (48 trials) were conducted per link, making a total of 96 trials. This word game was combined with maintenance training for tacting the aliens. When she was retested for Leg 2 one last time, the results were exactly the same as before: complete reversal (0/48 trials correct).

The third and last child to fail the test for Leg 2, was Louise. After the test she had four extra sessions. Tacting and the intraverbal word games were still in place. The new "You say what I say" game with extension was then played with her. Experimenter 1 said: "lup-noom-lup-noom ... a lup is a noom" and "tor-zaag-tor-zaag ... a tor is a zaag" which Louise echoed. Four 8-trial blocks (32 trials) were conducted per link, making a total of 64 trials. In the retest that followed, not much improvement was found (11/12 correct for Block 1, but then 5/12, 6/12 and 4/12. After Block 1, Louise mostly selected the animal on the left. When retested again for Block 2, 3, and 4 two days later, the scores were similar (4/12, 5/12 and 6/12 correct), and Experimenter 2 commented that Louise was hardly even looking at the animals when prompted and when she selected one of them as zaag or noom).

On the repeated trials with why-questions after the test, Louise was presented with Lup 1 and Tor 1. She was asked for the noom, and selected correctly. To the why-question she responded, "It's got spikes". In the next trial, she picked the zaag correctly and said, "It's got points". Then, she was presented with a different pair of aliens (Lup 2 and Tor 2). Louise picked the incorrect animal when asked for the noom. She said that the stimulus she had selected was the noom, because "It's got points on its head". When asked for the zaag, she also picked incorrectly, commenting: "It's got points at the end".

Discussion

The aim of Study 2b, and of 1b, was to investigate whether young children can learn higher-level names for objects simply as a result of learning intraverbal relations between lower-level names and potential higher-level names. In an echoic and intraverbal word game, each of the lower-level names hib, feb, tor, and lup were linked with a higher-level name, either zaag or noom. The word game was played without the aliens being present. Following this, the children were tested for listener behaviour with regard to the potential higher-level names (i.e., hear /higher-level name/ – select animal).

The data show that of the 8 children who completed training and were tested, seven passed the listener behaviour test at the higher name level for Leg 1. Lee did not pass initially, but he did after several retests (see results, for details). All 7 children who passed the test for Leg 1 the first time, were then given the relevant intraverbal training and tested for Leg 2, and 3 of them also passed that test. So in the case of Leg 1, for all 8 children it was sufficient to learn the intraverbal relation between lower and higher-level names to also acquire listener behaviour at the higher name level, without direct training. This adds to the data of Study 1b, and is in line with Horne and Lowe's (1996) description of interchange of listener behaviour across intraverbally linked names (see the discussion of Study 1b, in Chapter 3). In the test for Leg 1, 2 children (Adam and Simon) uttered lower-level names once or twice during the test, and one of them (Simon) produced an intraverbal name relation after picking an alien in one of the trials. There were no vocalizations during the test for Leg 2. After the tests, only Simon explained his behaviour by referring to the lower-level names; the other children either said "Don't know" or referred to physical features of the animals.

There were two puzzling phenomena in the data. First of all, it was surprising that of the 7 children who passed the test for Leg 1, 4 failed the test for Leg 2, even though all the component skills (lower-level naming, and the intraverbal relations) were in place. The second surprise was that 3 children who failed the test for Leg 2 showed reversal of responses for some of the aliens (Ginny and Louise), or even full reversal for all four animals (Tomos).[3] When retested, Ginny showed full reversal for all four animals as well.

After failing the test for Leg 2, these 3 children had extra sessions with a further word game that gave them more explicit rules (e.g., a hib *is a* zaag). As was the case in Study 1b, the data show that this did not help them to pass the test at later stages. Full reversal of responses was found for Ginny and Tomos, also in retests after the additional sessions.

Study 2b was the same as 1b, apart from the fact that pre-training with the names of the familiar stimuli and testing for the corresponding listener behaviour was added. This pre-training was designed to provide a real life parallel, and to ensure that the child would respond appropriately to the procedure and the instructions. And Study 2b thereby had the additional aim of investigating whether *more* children would show correct listener behaviour at the higher name level with regard to the aliens, when pre-training and pre-testing was added. When comparing the data of Studies 1b and 2b (which will be done in further detail in the general discussion in Chapter 5), after pre-training more children passed the test with the alien animals for Leg 1 (7/8, plus 1/8 after a retest), than when children had no pre-training (5/8). (The sample is too small to determine whether this difference is due to chance.) However, for Leg 2, in each of the groups, 3 children passed the test.

---

[3] In Study 1b, reversal of responses was also found in Lyn's test data for Leg 2.

Study 2c – Names and transfer of behaviour

As Horne & Lowe (1996) described, listener behaviour (or functional responding) with regard to objects can transfer without direct training to other objects that are members of the same name relation (recall the example of listener behaviour appropriate to either a boat or a hat, which could transfer to a plastic container once the child is told the container is a boat or a hat, respectively — see the introduction of Study 1c). This is called *transfer of function*. Now, would a child also show transfer of function at different name levels (the lower and the higher level), once she has learned names at different levels for particular objects? That was the focus of Study 2c, as it was in 1c.

In Study 1c, two functional responses (gestures) were trained to two aliens (i.e., see alien – produce gesture), one to a zaag, one to a noom. After training, the children were tested for transfer of function to the other stimuli by presenting the child with the aliens not used in training, and the question "How does this one go?" Study 2c was exactly as Study 1c, except that in the former study pre-training with familiar stimuli preceded the procedure with the aliens. In pre-training, both the stimuli and the gestures were different, but otherwise the procedure was exactly as in Study 1c; two gestures were trained, one to an animal, one to a food. After training, the child was tested for transfer of the two functions to the other familiar stimuli (i.e., see stimulus– produce gesture).

The aim of Study 2c was to see whether providing a procedural parallel with familiar stimuli would lead to more children showing transfer of function with the aliens, and whether transfer would extend more consistently to both levels (in Study 1c, only 1 child showed full transfer of function, while the other 2 showed partial transfer of function at both levels). The added pre-training aimed to ensure that the children would respond appropriately to the instructions, as well as providing familiarity with the procedure employed.

Method

Participants

Of the participants in Study 2b who passed the test for listener behaviour at the higher level for all legs of the hierarchical structure, 3 children (all male) were available to take part in Study 2c. They were between the ages of 3 years and 6 months, and 4 years and 6 months at the start of the study (see Table 3.6).

Table 3.6 Participants' sex, age at start of pre-training with familiar stimuli, age at start of alien training, and age at first test with the aliens.

Participant	Sex	Age at start of pre-training (years/months*)	Age at start of alien training (years/months*)	Age at testing (years/months*)
Huw	M	3/6	3/6	3/6
Mike	M	4/0	4/1	4/2
Simon	M	4/6	4/6	4/6

* age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).

Apparatus and Stimuli

Apparatus and stimuli in Study 2c, and all the following studies, were as in Study 1a. In pre-training, and in testing at the end of pre-training, the eight familiar stimuli were used (see Study 2a).

Procedure

As before, a single case design with replication across participants was employed. No parents were present during the study. There were one or two daily sessions with the child, which took about 10 to 20 min each. For each participant, Study 2c was started as soon as possible after completion of Study 2b.

Novel function training to familiar stimuli. For each child two of the eight familiar stimuli, one animal and one food, were randomly selected for novel function training. The two novel responses to be trained, were two gestures, the first one to be trained to an animal, the second one to a food. The gestures were: 1) hand moving forward in a "stroking" movement (brief and fast movement, fingertips pointing forward while palm is facing down), and 2) hand to mouth (as in bringing a piece of food to the mouth). The gestures are shown in Figure 3.9.

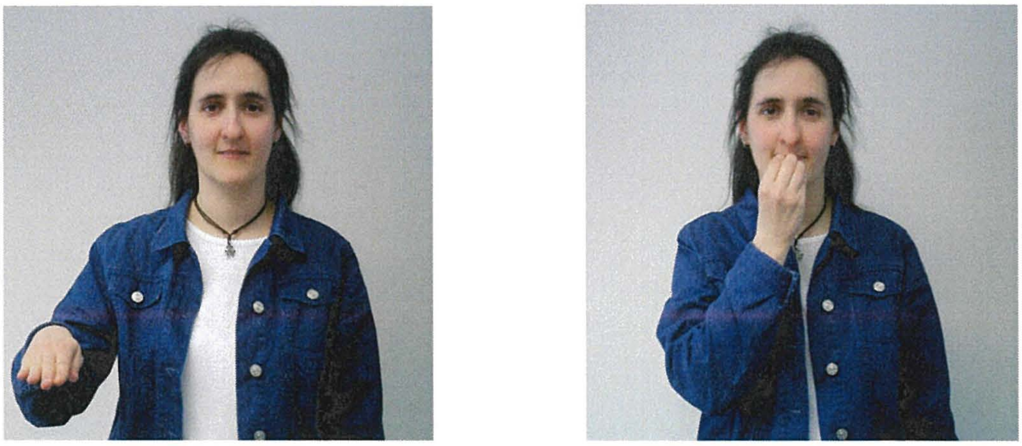


Figure 3.9 The functional responses, trained to one animal stimulus (left), and to one food stimulus (right).

During training the two familiar stimuli were placed in front of the child. In the first trial for each shape, the experimenter pointed at the target stimulus and introduced it by saying: “Look, this one goes like this (experimenter demonstrates gesture). Can you do this?” In subsequent trials with the same shape, the experimenter pointed at the targeted stimulus and said, “How does this one go?” A correct response was followed by social praise (e.g., "Yes, well done!")[4], and incorrect responses by corrective feedback: "It goes like this. Can you do this?" Apart from the stimuli and the gestures, all procedural features were exactly as in Study 1c. During the function-training period, maintenance tact training was conducted with the familiar stimuli (Set 1 and 2), and the intraverbal word game.

Reduction in reinforcement rate. As before, in preparation for the test situation reinforcement was reduced to 0% in the last training stage. In these no reinforcement sessions the experimenter used the penguin hand puppet (Peter Penguin), and asked the child to teach Peter the game. At this stage, as before, the learning criterion was seven out of eight correct within a block over two consecutive blocks. If the performance of any child did not meet the zero reinforcement criterion, then reinforcement was reintroduced until responding was once again at criterion. The child was then tested once again in extinction, and so on, until the child's performance met the zero reinforcement criterion.

Category transfer-of-function test. Following novel function training, the child was tested for transfer (i.e., see stimulus– make gesture) of these responses to the other familiar stimuli. At this stage, again, all details of the procedure were the same as in Study 1c. After the review trials, Experimenter 1 conducted the transfer test with the familiar stimuli. Two familiar stimuli (a random selection of one animal and one food) were placed on the table in pre-specified positions. The experimenter pointed at one of the two stimuli, and asked: "How does this one go?" No reinforcement was given during the test.

---

[4] And occasionally, in addition, a music book was presented on which the child could press a button.

Novel function training to arbitrary stimuli, and testing for transfer of function. Following pre-training and testing with the familiar stimuli, the complete procedure of Study 1c was implemented.

Again, each response was randomly assigned to a particular animal for each child. The relations between the specific animals, their names at both levels, and the trained responses for each participant are shown in Table 3.7.

Table 3.7 For each child, the novel behaviour trained to the specified stimulus (See Study 1c, for photos of the gestures).

Participant	Gesture	
	Hand on shoulder	Hands in front
Huw	 Tor 2 / zaag	 Lup 1 / noom
Mike	 Hib 2 / zaag	 Tor 1 / noom
Simon	 Tor 1 / noom	 Feb 2 / zaag

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 100%. In addition, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Pre-training with familiar stimuli: function training and reinforcement reduction. All 3 participants completed the pre-training with familiar stimuli. None of them needed more than the minimum number of 40 trials; 24 trials for the function training with reinforcement, and 16 for function training without reinforcement[5].

Transfer of function test with familiar stimuli. The children were all tested for transfer of the trained functions to the other familiar stimuli not involved in training. All passed the test. The performance of Mike and Simon was errorless over 60 trials. Huw made two errors.

Arbitrary stimuli: function training and reinforcement reduction. All 3 children then completed the function training with the alien animals. They needed on average 56 trials (range 40—88); 37 trials for the function training with reinforcement (range 24—64), and 19 for function training without reinforcement (range 16—24). Mike and Simon needed no more than the minimum number of trials for both stages (24 for training with reinforcement, and 16 for training without), while Huw had more training trials with reinforcement, in part because he had difficulty producing the *hands in front* gesture with the palms facing down (eventually his version of the gesture, with the palms facing him, was accepted), and in part because he started to tease by making the incorrect gesture … with a smile!

Arbitrary stimuli transfer of function test. Table 3.8 shows the alien pairs that were used in training and testing, and the number of correct responses out of 10 per pair for each child. The transfer of function criterion was 80% (8 out of 10) correct

---

[5] Due to experimenter error, Huw was not given trials without reinforcement.

for each of the two trial blocks per alien pair. Simon was the only one whose performance met the transfer of function criterion for all three pairs. Huw reached criterion for Pair 2 only, and Mike reached it for none of the pairs. The other criterion was 80% percent (8 out of 10) correct for each alien.

Table 3.8 The alien pairs used in training and testing, and the test scores for the three alien pairs (P1, P2, P3), for each child. The transfer of function criterion was 80% (8 out of 10) correct for each of the two 10-trial blocks per alien pair. The top score for each pair shows the result for the first test trial block, the second score gives the result for the second trial block.

Child	Gesture	Training stimuli	Test stimuli Pair 1	Pair 2	Pair 3	Scores per pair (out of 10)
	Hand on shoulder					P1: 5 5
Huw		Tor 2/zaag	Feb 1/zaag	Tor 1/zaag	Feb 2/zaag	P2:10 10
	Hands in front					P3: 4 0
		Lup 1/noom	Lup 2/noom	Hib 1/noom	Hib 2/noom	
	Hand on shoulder					P1: 5 0
Mike		Hib 2/zaag	Feb 2/zaag	Feb 1/zaag	Hib 1/zaag	P2: 6 10
	Hands in front					P3: 5 1
		Tor 1/noom	Lup 1/noom	Tor 2/noom	Lup 2/noom	
	Hand on shoulder					P1: 10 10
Simon		Tor 1/noom	Hib 1/noom	Hib 2/noom	Tor 2/noom	P2: 10 10
	Hands in front					P3: 9 10
		Feb 2/zaag	Feb 1/zaag	Lup 1/zaag	Lup 2/zaag	

Figure 3.10 provides the percentage of correct responses per child to each of the six aliens presented in the test (i.e., the aliens not involved in training). Two functions had been trained, one to a zaag (e.g., Tor 2, for Huw), one to a noom (e.g., Lup 1, for Huw). Per child it can be seen in Figure 3.10 whether the functions transferred without direct training to the other members of the lower-level categories (for Huw, Tor 1 and Lup 2), and whether they transferred to the other members of the higher-level categories (the other zaags and nooms, in Huw's case, the hibs and the febs). Like Jon in Study 1c, Simon showed almost perfect transfer of function both at the lower level and the higher level; he made just one error over 60 trials (10 trials for each of the six aliens). Huw showed transfer to three aliens: to both the aliens belonging to the same lower-level category as the aliens used in training, and to one alien belonging to the same higher-level category as one of the aliens involved in function training. For two aliens Huw showed reversal of responses, and a strong tendency to reversal for the remaining alien. Mike showed transfer of function to two aliens, one belonging to the same lower-level category as one of the aliens used in training, and one belonging to the same higher-level category as one of the aliens involved in function training. He showed reversal of responses for one alien, and a tendency towards reversal for three other aliens.

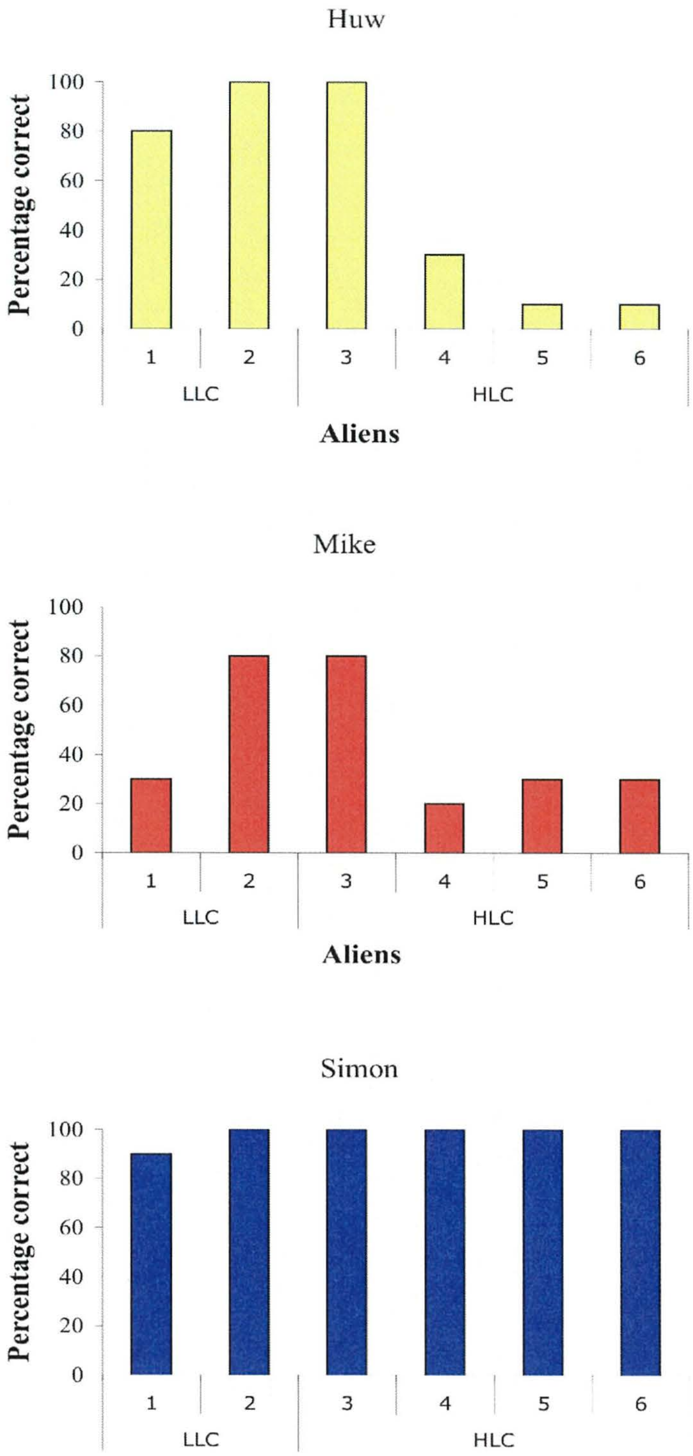


Figure 3.10 The percentage of correct responses for all six aliens used in transfer of function testing, two of the same lower level category (LLC), and four of the same higher level category (HLC) as the aliens used in function training. The second transfer of function criterion was 80% (8 out of 10) correct for each alien.

The children made no vocalisations during testing. As before, one or more trials were repeated in a posttest interview. The child was asked: "Why does it go like that?" Huw produced the incorrect gesture for Hib 1, and in reaction to the why question he said “Don’t know ... hib”. Experimenter 1 asked “And does a hib go like that?” Huw confirmed that this was the case. He then reacted in the same way for Tor 1 (for which he did produce the correct gesture). Mike referred to physical features. He produced the correct gestures for Tor 2 and Feb 1, and explained his choice of gesture by saying for Tor 2 (a noom) “ 'Cause it’s straight”, and for Feb 1 (a zaag) “ 'Cause it’s pointy”. He produced the incorrect gestures for the other four aliens, and explained by saying for Feb 2 (a zaag) “Because it’s straight”, for Lup 1 (a noom) “Cause he’s pointy up here”, for Lup 2 (a noom) “Cause this horn is pointing here”, and for Hib 1 (a zaag) “Cause these horns are pointing here”. So all in all, his explanations were not consistent, either within each of the lower level categories, or within each of the higher level categories. Simon produced correct gestures on all trials but reacted to the why-questions with “Don’t know”.

Additional sessions. Both Huw and Mike were retested, Huw for four trial blocks during which his attention was poor, Mike for all six trial blocks. Their responses in the retest were quite consistent with their original test responses, and where there were changes, they made no difference to the transfer of function data per alien. For Huw, if the scores for the repeated sheets are combined with the scores for the two sheets that were not repeated, then his scores for Pair 1 would be 5/10 and 10/10, while for Pair 2 they would be 10/10 and 10/10, and for Pair 3 the scores would be 0/10 and 0/10. For the individual aliens, he scored 90% instead of 80% for Alien 1, he showed no change for Alien 2 and 3 (both 100%), and for Alien 4 and 6 his score decreased from 30% and 10%, respectively, to 0%. The only more substantial improvement was that for Alien 5 his performance improved from 10% correct to 60%. But again, this made no difference to the transfer of function data.

Mike's responses in the retest showed a more coherent pattern, compared to his original test scores. For Pair 1 he scored 0/10 and 0/10, for Pair 2 it was 10/10 and 10/10, while for Pair 3 it was 0/10 and 2/10. At the level of the individual aliens, for Alien 1 his score fell from 30% to 20%, for Alien 2 and 3, his score increased from 80% to 100%, and for Alien 4, 5, and 6 his scores decreased to zero from 20%, 30% and 30%, respectively. At the end of the retests he showed perfect transfer of function for Alien 2 and 3, and reversal of responses for the other four aliens.

Discussion

The aim of Study 2c, and of 1c, was to investigate name-based transfer of function. Would responses (gestures, comparable to greetings) taught to one member of one higher-level category (one zaag) and one member of the other higher-level category (one noom) transfer to the aliens that were not involved in training? And if so, would there be transfer just at the lower name level, or also at the higher name level?

The data show full transfer of function at both levels for only 1 (Simon) of the 3 participants. Huw reached criterion for one pair of aliens, whereas Mike did not reach criterion for any of the pairs. Their percentage of correct responses at the level of the individual aliens did suggest that Huw showed transfer of function for three aliens, the two aliens of the same lower-level category as the aliens involved in training, and one alien of the same higher-level category as one of them. Mike's data at the level of the individual aliens suggested that he showed transfer at the lower level, for one alien, and at the higher level, for another alien. So although he showed transfer of function at both levels, transfer was consistent at neither of the two levels, which was also the case for Alun in Study 1c. The two aliens for which Mike reached criterion had been presented together. Therefore, as was noted for Alun before, it could be the case that his transfer at the higher name level was brought about through transfer of function at the lower name level, in combination with "selection by exclusion" (see the discussion of Study 1c, in Chapter 3). But the question that was asked for Alun can also be asked for Mike: why does his behaviour not transfer to the other alien in the same lower-level category as the aliens involved in training? In Mike's case, this might be due to him basing his responses on shape rather than names, as his reactions to the why-questions that followed the test suggest. Only Huw responded to the why-questions by referring to (the lower-level) names, but he did not produce correct gestures for all aliens. Interestingly, reversal of responses was found again, as it was in Study 2b. Mike reversed responses for Alien 4 (20% correct), and showed a tendency to do the same for Aliens 1, 5 and 6 (30% correct, each). In the retest, similar behaviour was observed; this time he showed full reversal for them (20% correct for Alien 1, 0% for Aliens 4, 5, and 6).

Transfer of function as shown by Simon and Huw is in line with Horne and Lowe's (1996, p.213) description of how functional responding may transfer by means of names (see the introduction for Study 1c). As pointed out in the discussion of Study 1c, we know that the children can name at the lower level because they reached criterion in tact training and they passed the listener behaviour test. But it is not clear whether the children can name at the higher level. They showed appropriate listener behaviour at the higher level, in the tests for Study 2b, but we don't know whether they have the corresponding speaker behaviour at the higher name level in their repertoire, because they were not trained or tested for it. So some children may have been able to tact the aliens at the higher name level, while others may not, and this may have influenced their performance in the transfer of function test.

A further aim of Study 2c was to investigate whether added pre-training, ensuring appropriate responding to instructions and familiarity with the procedure, would lead to more children showing transfer of function with the aliens, and whether there would be more consistent transfer to both the lower and the higher level. As in Study 1c, the data show full transfer of function for one child, and partial transfer of function at both levels for the two other participants. So the added pre-training did not seem to make a difference.

Study 2d – Names and transfer of a second behaviour

Study 2d replicated Study 1d. Both studies aimed to extend transfer of function data and to see whether transfer of behaviour would be consistent within the same child — that is, whether a child showed a regular behaviour pattern in testing. This time the trained novel responses were not gestures, but animal cries ("Boo!" and "Raagh!"). The cries were trained one to a zaag and one to a noom. After training, the child was tested for transfer of function to the other stimuli in the respective higher-level categories (see alien – produce cry).

Because there was no equivalent of the animal cries for the two higher-level categories of familiar stimuli (animals and foods), there was no pre-training in Study 2d. However, the pre-training the participants had received in Study 2c was also useful for them in Study 2d, because it had shown them the general procedure of training for two functional responses, followed by transfer of function testing.

Method

Participants

All 3 participants in Study 2c (all male) were available to take part in Study 2d. They were between the ages of 3 years and 8 months, and 4 years and 6 months at the start of the study (see Table 3.9).

Table 3.9 Participants' sex, age at start of alien training, and age at first test.

Participant	Sex	Age at start of alien training (years/months*)	Age at testing (years/months*)
Huw	M	3/8	3/8
Mike	M	4/2	4/2
Simon	M	4/6	4/6

* age in days rounded up to nearest whole month (i.e., 16 days and more rounded up).

Procedure

Novel function training to arbitrary stimuli, and testing for transfer of function. The complete procedure of Study 1d was implemented. The two responses ("Boo!" and "Raagh!") were assigned to particular animals for each individual child, one to a zaag, one to a noom. For each child, the two alien animals used in function training in Study 2d belonged to different lower-level categories than the two used for training in Study 2c (in the case of Simon, for example, in Study 2c the aliens employed in training were a tor and a feb, whereas in Study 2d a hib and a lup were used). Each response was randomly assigned to one animal. The relations between the specific animals, their names at both levels, and the trained responses for each participant are shown in Table 3.10.

Table 3.10 The novel behaviours, the cries "Boo!" and "Raagh!", were allocated to one zaag and one noom.

Participant	Cry	
	Boo!	Raagh!
Huw		
	Feb 1 / zaag	Hib 2 / noom
Mike		
	Lup 1 / noom	Feb 2 / zaag
Simon		
	Hib 2 / noom	Lup 1 / zaag

Interobserver reliability. An independent observer scored all trials in a randomly selected 25% of all training sessions; interobserver agreement on these trials was 100%. Similarly, all test trials were scored; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Function training and reinforcement reduction. There were 3 participants in Study 2d. They all reached criterion for novel behaviour training in the minimum of 40 trials; 24 trials for the training with reinforcement, and 16 for training without reinforcement.

Transfer of function test. Table 3.11 shows the alien pairs used in training and testing, and the scores per pair. The transfer of function criterion was 80% (8 out of 10) correct for each of the two trial blocks per alien pair. Huw and Simon reached this criterion for one trial block each, but for no pair do they reach it for both blocks. Mike reached it for none of the pairs either. The other criterion was 80% percent (8 out of 10) correct per block over two consecutive blocks, for each alien. Figure 3.11 shows the percentage of correct responses to each of the six aliens that had not been involved in training. As in Study 1c, two functions had been trained, one to a zaag (for Huw, for example, this was Feb 1), one to a noom (for Huw, Hib 2). Figure 3.11 shows whether there was any transfer of function within the lower-level categories (in Huw's case, transfer to the category members Feb 2 and Hib 1), and within the higher-level categories (that is, to the other zaags and nooms, in Huw's case the tors and the lups). Simon, who showed transfer of function to *all* aliens in Study 2c, reached criterion for one alien only, this time; transfer of function was found to an alien belonging to the same lower-level category as one of the aliens used in training. Huw and Mike reached criterion for none of the aliens. Instead, reversal of responses was found in the test results for all 3 participants. Mike showed complete reversal, getting zero correct in 60 trials. Huw's performance consists of reversal for four aliens, and Simon's for two. Although Mike shows reversal of responses, his responses are completely consistent. That is not the case for Huw and Simon. Accordingly, because of the difference between Simon's performance in Study 2c (full transfer of function) and Study 2d (transfer to only one alien), both Huw and Simon were retested, the results of which are described at the end of this section.

Table 3.11 The alien pairs used in novel function training, and in testing for transfer of function; test scores are presented for each the three stimulus pairs (P1, P2, P3). The transfer of function criterion was 80% (8 out of 10) correct for each of the two 10-trial blocks per alien pair. The top score for each pair shows the result for the first test trial block, the second score gives the result for the second trial block

Child	Cry	Training stimuli	Testing stimuli: Pair 1	Testing stimuli: Pair 2	Testing stimuli: Pair 3	Scores per pair (out of 10)
	Boo!					P1: 1 1
Huw		Feb 1/zaag	Tor 2/zaag	Tor 1/zaag	Feb 2/zaag	P2: 0 0
	Raagh!					P3: 1 8
		Hib 2/noom	Hib 1/noom	Lup 2/noom	Lup 1/noom	
	Boo!					P1: 0 0
Mike		Lup 1/noom	Lup 2/noom	Tor 2/noom	Tor 1/noom	P2: 0 0
	Raagh!					P3: 0 0
		Feb 2/zaag	Hib 1/zaag	Feb 1/zaag	Hib 2/zaag	
	Boo!					P1: 0 6
Simon		Hib 2/noom	Tor 1/noom	Tor 2/noom	Hib 1/noom	P2: 10 5
	Raagh!					P3: 0 5
		Lup 1/zaag	Feb 2/zaag	Lup 2/zaag	Feb 1/zaag	

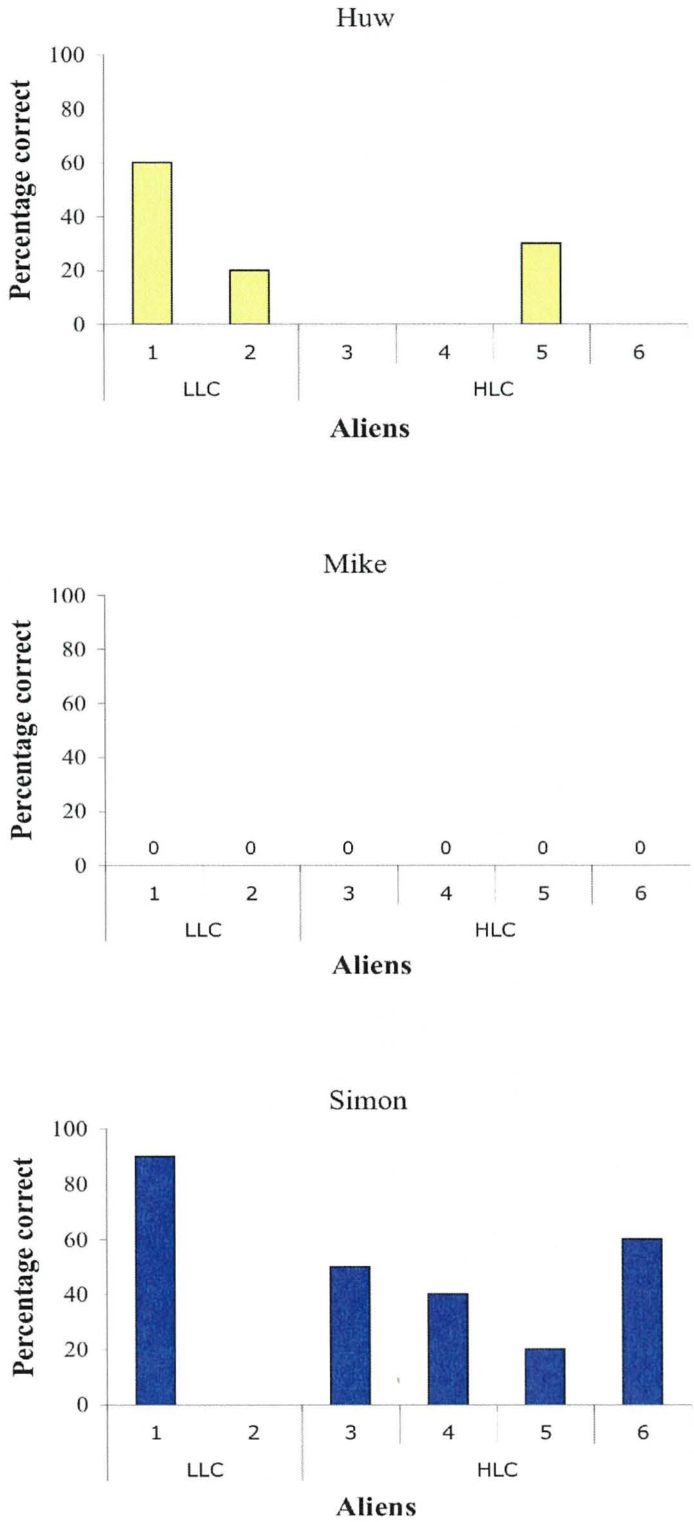


Figure 3.11 The percentage of correct responses for all six aliens used in transfer of function testing, two of the same lower level category (LLC), and four of the same higher level category (HLC) as the aliens used in function training. The second transfer of function criterion was 80% (8 out of 10) correct per alien.

The children's vocalisations during testing are listed in Table 3.12. As before, one or more trials were repeated after testing, and following a response the experimenter asked the child, "Why does it say that?" Huw consistently replied to this by giving the (correct) lower-level name of the alien " 'Cause it's a hib/feb/tor/lup". However, the cry he gave for the aliens was incorrect in four out of six cases. Mike referred to physical features, which he also did in Study 2c. He produced the correct cries for Feb 1, and Tor 2, and explained his choice of cry by saying for Feb 1 (a zaag) " 'Cause he got a horn", and for Tor 2 (a noom) " 'Cause he's got this (points at (front) end of alien)". In Study 2c, he said for this alien " 'Cause it's straight", as opposed to pointy, or with horns. He produced the incorrect cries for the other four aliens, and explained by saying for Hib 1 and 2 (zaags) " 'Cause he/it's got a tail", for Tor 1 (a noom) " 'Cause he's got this (points at alien's long and pointy nose)", and finally, for Lup 2 (a noom) " 'Cause it's got a horn." So again, his explanations were not consistent, either within each of the lower level categories, or within each of the higher level categories. Simon produced incorrect cries on the repeated trials and reacted to the why-questions with "Don't know".

Table 3.12 The children's vocalisations during testing for transfer of function (Test 2d).

Participant	Stimuli	Experimenter's prompt	Child's vocalisations (apart from the cries)
Huw	Feb 2 + Lup 1	What does this one say?	Feb (E: What does it say?) Raagh, no Boo (E: So what does it say?) Raagh.
Simon	Tor 1 + Feb 2		I need a clue. (E: Well, it can either say Boo! or Raagh!)
	Lup 2 + Tor 2		I need a clue. (E: Look at the animal, and think of all the games, that can help you. And just do what you think.)
			(Between trials:) It doesn't really matter what I say, does it? (E: It *does*, Simon, please think carefully!)

The underlined stimuli are the targeted stimuli.

Additional sessions. Huw and Simon had additional sessions. For Huw these involved maintenance training, two complete retests, and one partial retest that was discontinued because Huw reverted to Raagh! responses for all trials. In the first retest, Huw's scored for Pair 1 were 8/10 and 10/10, for Pair 2 they were 10/10 and 5/10, and for Pair 3 they were 6/10 and 4/10, meaning that he only reached criterion for Pair 1, the pair for which he had reversed the responses in the original test. In that test he reached criterion for none of the pairs. As can be seen in Figure 3.12, at the level of the individual aliens he showed transfer of function for four aliens now, rather than none. Transfer was found for one alien in the same lower-level category as the aliens used in training, and to three out of the four aliens that belong to the same higher-level category. During the second trial block for both Pair 2 and 3, Huw reverted to producing the Raagh! response in all 10 trials for Pair 2, and in 9 out of 10 trials for Pair 3.

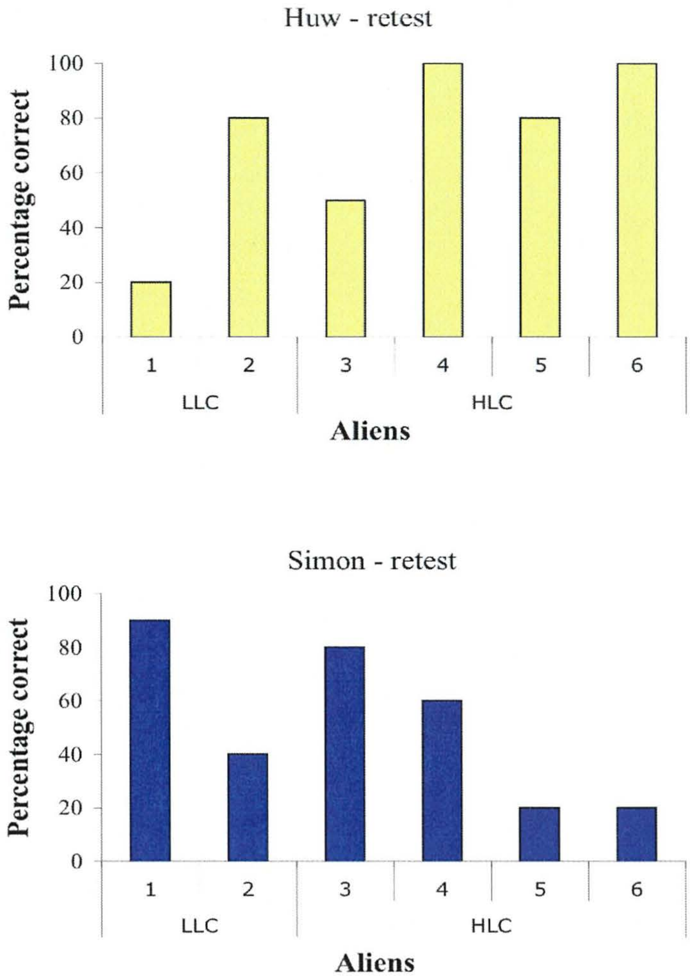


Figure 3.12 The percentage of correct responses for all six aliens used in retests for transfer of function, two of the same lower level category (LLC), and four of the same higher level category (HLC) as the aliens used in function training. The second transfer of function criterion was 80% (8 out of 10) correct per alien.

In the second retest, after more maintenance training, he reached criterion for none of the pairs, but at the level of the individual aliens he showed transfer of function to two aliens (Alien 2, which was Hib 1, and Alien 5, Lup 1), one in the same lower-level category as the aliens used in training, and one in the same higher-level category. Furthermore, he is one correct response away from showing transfer of function for another alien in the same higher-level category (Alien 6, Lup 2). Here too, Huw reverted to the Raagh! response in most of the trials; per block of 10 trials, the number of times he did this was 8, 10, 8, 9, 6, and 9, so in 50 out of 60 trials. Several times during this test, Experimenter 1 told Huw that one of the two aliens says "Boo!", and one says "Raagh!". Huw was tested several times, because he consistently performed well in all the component tasks (of Study 2a and 2b), and he responded to the why-questions afterwards by referring to (lower-level) alien names. Therefore, it was expected that Huw would do well in the transfer of function tests too. However, in the course of repeated testing for 2d, Huw looked more and more lethargic, and often responded with "Raagh!" over and over again, as described above. After the last retest, training was resumed, but when Huw showed the same lethargic look as in testing (which he had not done before, in training), and his performance dropped below criterion because he replied with "Raagh!" more than with "Boo!", the experimental sessions with Huw were discontinued.

Simon was retested the day after the original test. In the retest, he reached criterion for none of the pairs, as in the original test, and it can be seen in Figure 3.12 that he showed transfer of function to two aliens rather than one (Alien 1, Lup 2, as before, and this time also to Alien 3, Feb 1). One of these belonged to the same lower-level category as the aliens used in training, and the other one belonged to the same higher-level category. In the second part of Test 1, Simon reacted with the "Raagh!" response in 8, 8, and 10 trials per (10-trial) block, respectively. In the retest, in three trial blocks he gave the "Raagh!" response in 8, 9, and 8 trials, respectively.

There were no vocalizations by the children during the retest, and replies to the why-questions for repeated trials afterwards, were as before.

Discussion

As did Studies 1c, 1d, and 2c, the present study aimed to investigate whether functional responses (animal cries) taught to one zaag and one noom, would transfer to other aliens, within the corresponding lower-level or higher-level category, or both.

The data show hardly any transfer of function in the initial test, and partial transfer at both levels in the retest. None of the children reached criterion for any of the pairs in the initial test. At the level of the individual aliens, Simon did show transfer for one alien (at the lower level), Huw and Mike for none of them. Huw and Simon were retested. Huw then reached criterion for one pair, and at the level of individual aliens he showed transfer of function to four aliens, rather than one. Of these four, one was of the same lower-level category, and three were of the same higher-level category as one of the aliens involved in training. Simon reached criterion for none of the pairs again, but at the level of the individual aliens he showed transfer of function to two aliens, rather than one; this was one case of transfer at the lower level, and one of transfer at the higher level.

So although their performance was better than in the original test, Huw and Simon did not show consistent transfer of function to either of the two levels in the retest. It should be noted that Simon's transfer of function for the two aliens in the retest, was due to him replying "Raagh!" all the time. This was also the case for one of the three aliens that Huw showed transfer to, and it may have played a part in the transfer to another alien as well. For him, transfer to the other two aliens was not due to this (which was supported by the fact that he reached criterion for those two aliens as a pair as well). These two aliens, Hib 1 and Tor 2, were presented together, and therefore his transfer of function to Tor 2, as a case of transfer at the higher level, may have been due to transfer of function at the lower name level, in combination with "selection by exclusion" (as explained in Study 1c for Alun's results).

The phenomenon of reversal of responses was observed again, most strikingly in Mike's data. He showed full reversal of responses for all three alien pairs (he had zero correct responses over 60 trials). Because his responses were completely consistent he was not retested. Mike explained his responses by pointing to physical features of the animals. Huw reversed responses for four out of six aliens in the original test, and showed a tendency to reverse for another alien (30% correct). In the retest, Huw showed reversal for only one alien. Simon reversed responses for two aliens, both in the original test and the retest. Although Huw's performance was not at all flawless, he did explain his responses by referring to lower-level names. Simon could not explain his responses.

For better judgement of Simon's performance in the test and the retest, some details of the test situation are relevant. He was tested on a very hot day, one or two days before he was due to leave the nursery to go to school. He was tired, and not taking a long time to think about his responses. Probably because he didn't get any feedback on his performance, Simon said at some point during the test: "It doesn't really matter what I say, does it?" Experimenter 1: "It *does*, Simon, please think carefully!" All of this may, perhaps in part, explain the discrepancy between his test data in Study 2c (full transfer of function) and 2d (no transfer of function when the combined criteria apply).

The (partial) transfer of function in Huw's data is in line with Horne and Lowe (1996 – see Chapter 3, Study 1c).

Study 2e – Names, transfer of behaviour, verbal prompts and stimulus control

Study 2e replicated Study 1e. Both studies were a test only, focussing on the effects of different verbal prompts on the child's transfer of functional responses: specifically, would a verbal prompt like "What *can* this one *do*?" which is more general than the prompts used in Study 1c ("How does this one go?") and 1d ("What does this one say?"), have stimulus control over both types of previously taught functional responses (gestures and cries) in the absence of direct training? In other words, would the new prompt evoke both types of responses without direct training?

The test also aimed to extend the transfer of function data. Would the functional responses trained to four animals of different lower-level categories in Study 2c and 2d, transfer to other lower-level category members, or also to other higher-level category members?

There was no pre-training with familiar stimuli for Study 2e.

Method

Participants

One male participant (Mike) was available to take part in Study 2e. He was 4 years and 2 months old at testing.

Procedure

The complete procedure of Study 1e was implemented.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 100%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Transfer of function test. Study 2e was a further test for transfer of function also focussing on the verbal prompt and stimulus control. In Study 2c, Mike had learned to produce gestures (hand on shoulder, and hands in front, see Study 1c for a description and photos of the gestures) in response to the prompt "How does this one go?" In Study 2d, he had been taught to produce animal cries ("Boo!" and "Raagh!") in reaction to the prompt "What does this one say?"

Table 3.13 shows which of the aliens these behaviours were trained to in the two studies, for Mike. In Study 1e, a verbal prompt was employed which was more general than the ones in the previous two tests: "What *can* this one *do*?". Would this elicit both types of responses? Table 3.14 lists the order of presentation of the alien pairs.

Table 3.13 Overview of the functional responses and the animals they were trained to.

	Gesture		Cry	
Participant	Hand on shoulder	Hands in front	Boo!	Raagh!
Mike				
	Hib 2/zaag	Tor 1/noom	Lup 1/noom	Feb 2/zaag

Table 3.14 The alien pairs in Test 1e.

Session	Alien pair*	
1	Feb 2	Tor 2
	Hib 2	Lup 1
	Hib 1	Lup 2
	Feb 1	Tor 1
2	same as 1	

*One 10-trial block per pair.

The test was conducted in two sessions. In the review trials before the test, for the first time, Mike did not show errorless listener behaviour at the higher name level. When asked which one the zaag/noom was, he picked the wrong alien in two out of four trials. When these trials were repeated, Mike did pick the correct animal in one of the trials, and later also for the other trial[6]. Then Experimenter 2 took over the session and the test trials were started. In all eight trial blocks (10 trials per block), Mike produced animal cries only, to both aliens. So at no time did the more general verbal prompt have control over both types of responses (gestures and cries). The criterion was per pair 8/10 correct within a block, over two consecutive blocks.

With regard to animal cries, Mike reached this criterion for Pair 1 (Feb 2 and Tor 2: 10/10 and 9/10) but not for Pair 2 (Hib 2 and Lup 1: 0/10 and 0/10), nor for Pair 3 (Hib 1 and Lup 2: 0/10 and 0/10), or Pair 4 (Feb 1 and Tor 1: 5/10 and 0/10). In order to consider transfer of function data for the animal cries in this study, the scores per alien are depicted in Figure 3.13. Aliens 1-6 here are the same as Aliens 1-6 for Mike in Study 2d. The added aliens (7 and 8) were the two aliens involved in function training in Study 2d. The other transfer of function criterion was 80% correct per animal.

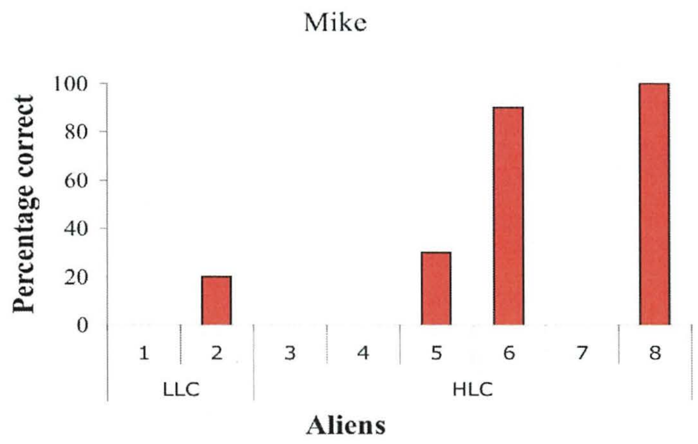


Figure 3.13 The percentage of correct responses for all eight aliens presented in Test 2e. Aliens 1-6 here are the same as Aliens 1-6 for Mike in Study 2d. Aliens 7 and 8 were involved in function training in Study 2d. The relevant transfer of function criterion was 80% (8 out of 10) correct per alien. The data presented here concern only animal cries. Mike produced no gestures during the test.

It seems that Mike showed transfer of function to two aliens for the animal cries, in Study 2e. However, one of the two aliens for which he scored higher than

---

[6] Three days later Mike's listener behaviour at the higher name level was checked again; 12 trials were conducted, all responses were correct.

80%, (Alien 8) had been involved in function training, and the other one (Alien 6) was the one it was presented with. Therefore, the high score for Alien 8 is not a case of transfer of function, and where it concerns Alien 6, one can hardly speak of transfer of function either. The other alien that had been involved in training in Study 2d, is Alien 7, which was presented together with Alien 4 in Test 2e. Strangely, Mike showed reversal of responses for both of these aliens. Because Mike's responses in Test 2e were limited to animal cries, the present data can be compared with the data from Study 2d. Mike showed complete reversal of responses for Aliens 1-6 in Study 2d. In Test 2e, he did this again for Aliens 1, 3, and 4. For Alien 2, he also reversed responses (in 8/10 trials), and for Alien 5, he showed a tendency to do the same.

Mike produced no vocalisations during the test, other than the animal cries. After the test, when why-questions were asked in repeated trials, Mike referred to physical features, as before. He produced the correct cries for Feb 2, and Tor 2, and explained his choice of cry by saying for Feb 2 (a zaag) " 'Cause he got horns here", and for Tor 2 (a noom) " 'Cause he's got a horn here". For the other six aliens, he produced the incorrect cries, and explained by saying for Hib 2 (a zaag) and Lup 2 (a noom) " 'Cause he's got a horn here", for Lup 1 (a noom) along with Hib 1 and Feb 1 (both zaags) " 'Cause he's got horns", and finally, for Tor 1 (a noom) " 'Cause he's got a nose and tail". As before, his explanations were not consistent, either within each of the lower level categories, or within each of the higher level categories.

Discussion

The aim of Study 2e was to see whether the new verbal prompt "What can this one do?", which was more general than the ones used in Studies 2c and 2d[7], would evoke both of the previously learned functional responses (gestures and cries) in the absence of direct training. And an additional aim of the study was to extend the transfer of function data, for both gestures and animal cries.

The data show that the more general prompt did *not* evoke both functional responses; Mike only responded by producing animal cries, not gestures. Therefore, this test was like a replication of the transfer of function test in Study 2d, and the data of both studies were comparable. However, this time the animals that the functional responses had been trained to were used in the test as well. They were each presented with an animal that had not been involved in training.

Mike reached criterion for only one of the four pairs of aliens, and that pair contained one of the animals that had been employed in function training for the animal cries. At the level of the individual aliens he also reached criterion for these two aliens only. Clearly, for the animal that had been employed in training the outcome cannot be classed as transfer of function. And Mike's correct responding for the other alien may have been due to selection by exclusion, rather than transfer of function by means of the names. Strangely, Mike did not reach criterion for the other alien animal that had been part of function training for the animal cries (Alien 7). He even showed full reversal of responses for that alien. Overall, he showed reversal for five of the animals (0% correct for four aliens, and 20% correct for the

---

[7] In Study 2c, the verbal prompt was "How does this one *go*?", and in Study 2d it was "What does this one *say*?".

fifth), and a tendency to do the same for the remaining alien (30% correct). Therefore, the data do not differ much from the data of Study 2d, where Mike showed full reversal for all six aliens (that had featured in training).

Apart from the animal cries, Mike produced no vocalisations during the test that could explain his behaviour. After the test, he explained his choices in terms of physical features of the animals, as he did before.

Horne and Lowe (1996) described how functional responding might transfer by means of names. In Mike's case this did not happen, because he based his responses on physical features rather than names.

Study 2f – Names, functional responses, listener behaviour, and transfer of function

Study 2f replicated Study 1f, but it did this in a more extensive form. Both studies were a test only, investigating whether without direct training children demonstrated the corresponding listener behaviour with regard to the functional responses – the gestures – taught in Study 2c (i.e., see gesture – select alien), when asked "Which one goes like this (E2 demonstrates gesture)?" Study 2f was more extensive than 1f: in testing, not only the aliens to which the gestures had been trained were presented, but also the other aliens. So in addition to being a test for listener behaviour, this was a test for transfer of function.

There was no pre-training with familiar stimuli for Study 2f.

Method

Participants

One male participant (Mike) was available to take part in Study 2f. He was 4 years and 2 months old at testing.

Procedure

Test for listener behaviour and transfer of function. The child was tested for listener behaviour (i.e., see gesture – select alien) and transfer of function with regard to the gestures trained in Study 2c. Whereas in testing in Study 1f, only the two alien animals to which the child was trained to emit the novel responses were used, this time all alien animals were used. Because Experimenter 2 needed to produce the gestures, it was not possible to use the test screen (see Study 1f). Each of the eight stimuli was targeted ten times over eight 10-trial blocks, making a total of 80 trials. Apart from this, all details of the procedure were the same as in Study 1f.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 99%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Test for listener behaviour and transfer of function. As in the previous study, in the review trials before the test, Mike did not show errorless listener behaviour at the higher name level. When asked which one the zaag/noom was, he picked the wrong alien in one out of four trials. When this trial was repeated, he again picked the wrong animal, but in another four trials he picked the correct animal. Then the other of the two aliens in that pair was targeted over six trials, and Mike picked the correct animal in all trials. At that point, Experimenter 2 took over the session, and Mike was tested for listener behaviour to the two functional responses (the gestures) he had learned to produce to two aliens in Study 2c (Tor 1 and Hib 2). Also tested was listener behaviour to the aliens that were *not* involved in training in 2c, which provided further transfer of function data. The test was conducted in two sessions. The listener behaviour criterion was per pair 80% (8 out of 10) correct per block, over two consecutive blocks. Mike reached this criterion for Pair 2 (Tor 2 and Feb 1: 9/10 and 10/10), and for Pair 3 (Tor 1 and Hib 2: 10/10 and 10/10). Pair 3 consisted of the aliens to which Mike had been trained to produce the gestures in Study 2c.

Mike reversed responses for Pair 1 (Lup 2 and Hib 1: 0/10 and 2/10), and for Pair 4 (Lup 1 and Feb 2: 0/10 and 0/10). The other listener behaviour and transfer of function criterion was 80% (8 out of 10) correct per alien. Figure 3.14 shows the percentage of correct responses per alien. Aliens 7 and 8 were part of training in Study 2c, and were presented together in testing. It can be seen that Mike's listener behaviour with regard to these aliens was perfectly in place. In the case of Aliens 2 and 3 (Tor 2 and Feb 1, presented together in the test), Mike's reaching criterion indicated transfer of function at the lower level, for Alien 2, and at the higher level, for Alien 3.

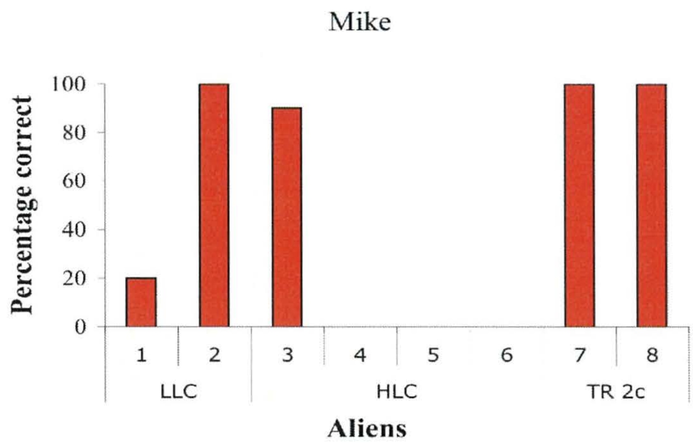


Figure 3.14 The percentage of correct responses for all eight aliens in the test for listener behaviour and transfer of function in Study 2f. Aliens 7 and 8 had been used in function training in Study 2c. Aliens 1 and 2 were of the same lower level category (LLC), while Aliens 3-6 belonged to the same higher level category (HLC) as the aliens used in this training. The relevant listener behaviour and transfer of function criterion was 80% (8 out of 10) correct for each alien.

Mike produced no vocalisations during testing. After the test, some trials were repeated, as usual, followed by why-questions. Mike referred to physical features, as he did after the test in Study 2c, and after all tests since.

For the combinations of Tor 1 and Hib 2, and Tor 2 and Feb 1, he picked the correct aliens again in response to both gestures, as in the test. And he explained his choice of alien for each gesture by saying for Hib 2 (a zaag), selected in response to the hand on shoulder gesture, " 'Cause he got a tail here", and for Tor 1 (a noom), selected in response to the hands in front gesture, "Together here straight (points at parts of the animal)". For Feb 1 (a zaag), selected in response to the hand on shoulder gesture, he says "horns", and for Tor 2 (a noom), selected in response to the hands in front gesture, "straight here". For the other four aliens, he picked the incorrect aliens in response to the gestures. He explained by saying "straight here" for Hib 1 and Feb 2 (both zaags which were incorrectly selected in response to the hands in front gesture); for Lup 1 and 2 (both nooms which were incorrectly selected in response to the hand on shoulder gesture), he said "horns here". For the first time, there is some consistency in his explanations, not connected to lower or higher-level names, because half of his responses are incorrect, but in the sense that for all the aliens he picked in response to the hands in front gesture he pointed out they they have a straight part, whereas for three out of four of the aliens he selected in response to the hand on shoulder gesture, he pointed to horns, and for the final one to a tail.

Discussion

As in Study 1f, the aim of Study 2f was to investigate whether children who had learned to produce two functional responses (gestures) with regard to one zaag and one noom, would demonstrate corresponding listener behaviour upon seeing the gesture (i.e., see gesture – select animal), without being directly trained to do so. Study 2f was more extensive than 1f. It tested not just for listener behaviour with regard to the aliens for which the gestures had been trained, but also the other aliens. Therefore, in addition to being a test for listener behaviour, Study 2f aimed to test for transfer of function.

The data show that listener behaviour with regard to the functional responses was perfectly in place for the aliens to which Mike had learned to *produce* the gestures. For the aliens that had not been employed in training, listener behaviour was in place for two out of six aliens (according to both criteria). Of these two, one belonged to the same lower-level category, while the other one belonged to the same higher-level category as one of the aliens involved in training. Given that the two aliens were presented together in the test, Mike's transfer of function at the higher level may have been due to transfer of function at the lower level, combined with selection by exclusion. For the other four alien animals, Mike reversed responses, even for the other alien belonging to the same lower-level category as one of the aliens involved in training. Again, he explained his behaviour referring to physical features of the aliens.

The emergent listener behaviour Mike showed is further evidence that some forms of behaviour can be established without the need for direct training.

Study 2g – Names, functional responses, listener behaviour, and transfer of function, once again

Study 2g replicated Study 1g, but it did this in the same more extensive form as Study 2f. Both Study 1g and 2g were a test only, investigating whether, without direct training, children would demonstrate the corresponding listener behaviour to the novel vocal responses, the animal cries, taught in Study 2d (i.e., hear /cry/ – select alien), when asked, "Which one says (E2 produces cry)?" Study 2g was more extensive than 1g, in that listener behaviour was tested with all the aliens and not just those that had featured in function training. So in addition to being a test for listener behaviour, this was a test for transfer of function.

There was no pre-training with familiar stimuli for Study 2g.

Method

Participants

One male participant (Mike) was available to take part in Study 2g. He was 4 years and 2 months old at testing.

Procedure

Test for listener behaviour and transfer of function. The child was tested for listener behaviour (i.e., hear /cry/ – select alien) and transfer of function with regard to the animal cries trained in Study 2d. While in testing in Study 1g, only the two alien animals to which the child was trained to produce the novel vocal responses (in

Study 1d) were employed, this time all alien animals were used. Each of the eight stimuli was targeted ten times over eight 10-trial blocks, making a total of 80 trials. Apart from this, all details of the procedure were the same as in Study 1g.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 100%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Test for listener behaviour and transfer of function. Unlike in the previous two studies, Mike gave correct responses in all review trials before the test, including the trials checking listener behaviour at the higher name level. Experimenter 2 then tested Mike for listener behaviour to the two vocal responses (the animal cries) learned with regard to two aliens in Study 2d (Lup 1 and Feb 2). Also tested was listener behaviour to the aliens that were *not* involved in training in 2d, which provided further transfer of function data with respect to the animal cries. The test was conducted in two sessions. The listener behaviour and transfer of function criterion was 80% (8 out of 10) correct per block, over two consecutive blocks. Mike reached this criterion for Pair 2 only (Lup 1 and Feb 2: 10/10 and 8/10). This was the pair of aliens with which Mike had been trained to produce the animal cries in Study 2d. For Pair 1 (Tor 2 and Hib 1), Pair 3 (Tor 1 and Feb 1) and Pair 4 (Lup 2 and Hib 2) he completely reversed responses (he scored 0/10 for all six trial blocks). The other criterion was 80% (8 out of 10) correct per alien. Figure 3.15 shows the percentage of correct responses per alien. Aliens 7 and 8 were part of training in Study 2d, and were presented together in testing. As in Study 2f, Mike's listener behaviour to these aliens was perfectly in place (that is, his behaviour was almost perfect – he made one error for each of the two aliens). For all other aliens, he showed complete reversal of responses.

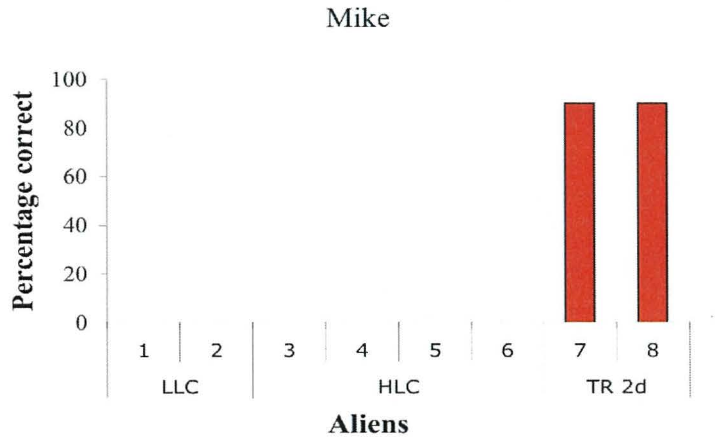


Figure 3.15 The percentage of correct responses for all eight aliens in the test for listener behaviour and transfer of function in Study 2g. Aliens 7 and 8 had been used in function training in Study 2d. Aliens 1 and 2 were of the same lower level category (LLC), while Aliens 3-6 belonged to the same higher level category (HLC) as the aliens used in this training. The relevant listener behaviour and transfer of function criterion was 80% (8 out of 10) correct for each alien.

Mike produced no vocalisations during testing. After the test, there were some repeated trials in a post-test interview. Mike pointed to physical features again, to explain his behaviour. For Lup 1 and Feb 2, he picked the correct aliens in response to both animal cries, as he did during the test. He explained his choice of alien for each animal cry by saying for Lup 1 (a noom selected in response to "Boo!"), "Cause straight here", and for Feb 2 (a zaag), selected in response to "Raagh!", "Cause he got horns here". For the other three pairs, he picked the incorrect aliens, in response to the animal cries. Here he explained by saying for Hib

1, Hib 2 and Feb 1 (all zaags), which he incorrectly selected in response to “Boo!”, “Cause he got horns (or: a horn) here”, and for Tor 2, Lup 1 and Lup 2 (all nooms), which he incorrectly selected in response to “Raagh!”, “straight here”. So for three pairs he was consistent in his explanations: animals that have one or more horns say “Boo!”, and animals that have a straight part say “Raagh!”. However, for one pair he explained it the other way around.

Discussion

As in Study 1g, the aim of Study 2g was to investigate whether children who had learned to produce two novel responses (animal cries), one to one of the zaags and another to one of the nooms, would demonstrate the corresponding listener behaviour upon hearing either one of these animal cries (i.e., hear /cry/ – select animal), without being directly trained to do so. Study 2g was more extensive than 1g. It was a test for listener behaviour not just with the aliens to which the animal cries had been trained, but also with the other aliens. So apart from being a test for listener behaviour, Study 2g also aimed to test for transfer of function.

The data show that listener behaviour to the novel vocal responses was in place for the aliens to which Mike had learned to *produce* the animal cries. He only made one error for each of them. For the alien animals that had not been involved in training, he reversed responses completely, for all six aliens, even for the two belonging to the same lower-level category as the aliens that the responses had been trained to. As before, he explained his behaviour by referring to physical features of the aliens.

Again, the emergent listener behaviour Mike showed, in Study 2f, and also here in Study 2g, is further evidence that behaviour can be established without direct training.

Study 2h – Category match-to-sample test

As mentioned in Chapter 3, Horne and Lowe (1996) proposed that categorization could be brought about without direct training, by learning the same name to a set of arbitrary stimuli (i.e., stimuli that share no characteristics apart from the name; so they have no physical features such as shape or colour in common, that distinguish them from other stimuli). This was tested both in Study 1h and 2h. Study 2h was a replication and an extension of 1h. Both studies consisted of a category match-to-sample test, which aimed to investigate whether children who had learned the names of the aliens could sort the alien animal stimuli (arbitrary stimuli, as defined above) into categories on the basis of the common name, without direct training. In the test, the child was presented with all eight alien animal stimuli. One of the aliens was selected to serve as the sample, and the child was then asked to give "the others" (see one category member – select the others). Would the child show category sorting at the lower name level, or also at the higher name level? If the child failed the category match-to-sample test, a second, altered version of the test was conducted, in which the child was prompted to tact (see alien – say "name") the sample stimulus at both levels before selecting the other category members.

Unlike Study 1h, in Study 2h pre-training with *familiar stimuli* preceded the procedure with the *aliens*. In pre-training, the child was trained to sort the familiar stimuli into categories, which was followed by a catgory match-to-sample test with the alien animal stimuli. The pre-training was added to provide a real life parallel and to ensure that the child would respond appropriately to the procedure and the instructions in the category test with the alien animals.

Method

Participants

One male participant (Mike) was available to take part in Study 2h. He was 4 years and 2 months old at the start of pre-training and the test with the familiar stimuli, as well as at testing with the aliens.

Apparatus and Stimuli

Apparatus and stimuli in Study 1h were exactly as in Study 1a. In pre-training the familiar stimuli were used (see Study 2a).

Procedure

Category match-to-sample training with familiar stimuli. The training with familiar stimuli was implemented in three phases. The difference between the phases was in the instruction the child was given. As a whole, this training could be seen as shaping the child's behaviour to respond correctly (on the basis of common names) to the *look-at-sample match-to-others* instruction, which is used in the category match-to-sample test 1. Throughout training, the screen that was normally used in test sessions, was on the table, between the experimenter and the child.

In Phase 1, the *tact-sample match-to-sample-tact* instruction was employed. On every trial, the experimenter put the eight familiar stimuli on the table in pre-specified positions in front of the child. Randomly, the experimenter picked up one of the stimuli and asked the child, "What's this?" When the child responded with the lower-level name (e.g., vegetable), the experimenter prompted the child to also provide the higher-level name (food) by asking, "And it's a …?" If the child did not respond, the latter question was repeated, but if the child still gave no response or responded incorrectly, the experimenter provided the correct response. If the child responded correctly, the experimenter asked, "Can you give me the other foods [animals]?" Correct sorting (of all three other higher-level category members) was followed by verbal praise. Following incorrect sorting, corrective verbal feedback was provided by pointing at the sample: "No, this is a food [animal], and this (pointing at the incorrectly selected stimulus) is an animal [food]." If the child only picked up one correct category member, for example the other lower-level category member, the experimenter asked, "Are there any more?" If the child then picked up one or both of the correct stimuli, the question was repeated until the child said "No". To avoid cueing the child for the correct number of stimuli to be selected, the question was repeated after each further stimulus selection the child made. Verbal praise was provided in case of a correct selection. When the child made correct selections on three consecutive trials per higher-level category, Phase 2 started.

In Phase 2, the *tact-sample match-to-others* instruction was employed. The procedure was as in the previous phase, but this time the target category was not specified in the instruction. On each trial, the experimenter selected the sample stimulus and asked, "What's this? Can you give me *the others* like this one?" As before, the criterion to move on to Phase 3 was correct selections on three consecutive trials per higher-level category.

In Phase 3, the *look-at-sample match-to-others* instruction was used. Again, the procedure was as in the previous phases, except that this time the child was not asked to tact the sample before making a selection. The verbal prompt was, "Look at this one. Can you give me the others like this one?" The categorization criterion was as before.

Category match-to-sample Test 1 with arbitrary stimuli. Before the test, Experimenter 1 conducted review trials with familiar stimuli and alien animals. The review trials consisted first of all, of eight alien animal tact trials, eight familiar stimulus word game trials, and eight alien animal word game trials. Furthermore, for checking listener behaviour with regard to the higher-level names, there was one trial for each of the four randomly selected pairs of one food and one animal of the familiar stimuli, and there was one trial for each of the four randomly selected pairs of one zaag and one noom of the alien animals. And finally, there were four trials to check the child's ability for category matching-to-sample with the familiar stimuli. In these trials all eight familiar stimuli were presented, and four of them were targeted, each by holding up the stimulus and saying "Look at this one. Can you give me the others like this one?" The criterion was correct selections on all four trials.

After the review trials, Experimenter 1 gave the child a general instuction at the start of the test: "You've played that game really well, with the yellow shapes[8]. Now we'll play it with the red shapes. We won't say 'Yes, well done' every time. We just want to see what you think, ok? So just do what you think." Then, Experimenter 2 conducted the category matching-to-sample test with the alien animals (i.e., see one category member – select the others). At this point, the complete procedure of Study 1h with the alien animals was implemented. A child who failed this test would be given category match-to-sample Test 2.

Category match-to-sample Test 2 with arbitrary stimuli. Before Test 2, Experimenter 1 conducted six practice trials with the familiar stimuli. In these trials a *tact-sample match-to-others* instruction was employed, involving tacting at both

[8] The yellow shapes are the familiar stimuli, the red shapes are the alien animals.

levels, as in Phase 2 of the category match-to-sample training with familiar stimuli that preceded Test 1. The categorization criterion was correct sorting on all six trials.

Category match-to-sample Test 2 with the alien animals was essentially the same as Test 1, apart from the fact that in the test trials, as in the review trials, a *tact-sample match-to-others* instruction was employed. In this case, the child was asked to tact the sample alien animal at the lower and higher level before selecting the stimuli from the same category. When presenting all stimuli and picking up one of them, Experimenter 2 would first ask, "What's this?" and then, "And it's a …?" If the child did not respond to this second question, a more specific question was asked, "And is it a zaag or a noom?" Finally, the child was asked, "Can you give me the others?" As mentioned, apart from the instruction, all details of Test 2 were the same as in Test 1.

Interobserver reliability. An independent observer scored all test trials; interobserver agreement on these trials was 96%. The independent observer reported no discrepancies between the scheduled and implemented procedures.

Results

Category match-to-sample training with familiar stimuli. In Phase 1, the *tact-sample match-to-sample-tact* instruction was used. The criterion was correct selections on three consecutive trials per higher-level category. Mike needed eight trials to reach criterion. In Phase 2, the *tact-sample match-to-others* instruction was employed. As in the previous phase, the criterion was correct selections on three consecutive trials per higher-level category. Mike reached criterion in the minimum of six trials. In Trial 1, when asked "Can you give me the others like this one?" rather than "Can you give me the other foods/animals?", Mike asked "You mean 'animals'?". Because this was pretraining, Experimenter 1 confirmed. In a later trial for this phase, when asked "Can you give me the others like this one?" Mike said "food" before making a selection. In the final trial, he asked "food?" and Experimenter 1 confirmed again. In Phase 3, the *look-at-sample match-to-others* instruction was used. The categorization criterion was as before. Again, Mike needed no more than the minimum of six trials.

Category match-to-sample Test 1 with arbitrary stimuli. Before the test, Experimenter 1 conducted review trials with familiar stimuli and alien animals. Mike made one error in the trials checking listener behaviour at the higher name level. He made one error with regard to a noom, and because of that, three more trials were added, two of which targeted that particular noom, and one that targeted a zaag. Mike responded correctly on all three trials. Therefore, Experimenter 2 took over, and tested the child for category matching-to-sample with the alien animals. Mike produced no correct category sorts on 48 trials. On one trial, he picked the correct three aliens, but added an incorrect one to it.

In general, rather than three stimuli, Mike selected between one and five stimuli per trial. In 29 trials, he did pick the other member of the lower-level category the sample alien belonged to, thereby reaching criterion for category sorting at the lower level. So, for example, when Tor 1 was targeted, he would select Tor 2. In seven trials (five of which were Feb 2 sample trials), he correctly selected the other members of the higher-level category the sample alien belonged to, but not the other lower-level category member. Overall, Mike was not consistent in the particular aliens, and the number of aliens, he selected for one specific sample, although he did show more consistency over the last trial blocks. Despite all this, Mike attended well to the task. When he had made a selection and was asked "Any more?" he usually said "no", with a firm tone. Mike's vocalisations during the test are listed in Table 3.15. After the last trial, Mike was asked, "why did you pick those?" Mike replied: "Animals, hippo, hippo, two hippo's, and two with horns". When Experimenter 1 asked: "Why are they like that one (pointing to the sample)?" Mike responded: " 'Cause they are animals". The same trial was then repeated, to see whether he would pick the same animals, which he did, while saying "two hippo's, two lups with horns on their back". When Experimenter 1 asked why, Mike said "'Cause animals". Experimenter 1 then asked: "Why are they like that one?" to which Mike replied: " 'Cause they say Raagh!" Following this, the pre-final trial of the test was repeated. Mike picked the same animals as during the test (only two aliens), but when asked why, he said he didn't know. Finally, one more test trial was repeated. Mike selected all the same aliens but one. When asked why, he said "two hippo's, two lups". Experimenter 1 then asked, "Why are they like that one?" and Mike responded for Lup 1 and 2, "Cause they got horns and tails", and for Hib 1 and 2 "horns".

Table 3.15 Mike's vocalisations during the category match to sample test.

Participant	Stimuli	Experimenter's prompt	Child's vocalisations
Mike	Feb 1 (zaag) +all others	Look at this one. Can you give me the others like this one?	feb (then correct selection of Feb 2, along with incorrect selections)
	Lup 1 (noom) +all others		I take four again (selected four aliens)
	Hib 1 (zaag) +all others		hippo (when picking up Hib 2)
	Lup 1 (noom) +all others		(counts three animals, picks up another and says:) four
	Tor 1 (noom) +all others		one, two, three, four (then puts in another animal, counts and takes it out of the hand of E2 again)
	Lup 1 (noom) +all others		(incorrectly picks up Hib 1 and 2, says:) hippo (Puts four animals in the hand of E2, picks up Feb 2, counts animals in E2's hand, takes Tor 2 out, and puts Feb 2 in instead.)
	Hib 2 (zaag) +all others		hippo (then selects Hib 1)
	Hib 2 (zaag) +all others		animal three (when he puts three animals on E2's hand)
	Hib 1 (zaag) +all others		hippy, hippy, hippy ... (correctly picks up Hib 2)
	Feb 2 (zaag) +all others		four animals ... three left (E1: what's that Mike?) three left on the table

Because Mike failed the category match-to-sample Test 1, he was given category match-to-sample Test 2. A few days after Test 1, and before Test 2, Mike was given maintenance training for tacting at the lower level, and for producing the intraverbals, and both were still in place. However, along with this, his listener behaviour at the higher name level was checked, and he did not do very well on this. In four trials, Mike made one error, so more trials were added, with corrective feedback where necessary. In the next four trials, he also made one error, in a further four trials he made two errors, and in another four trials he again made two errors. After one of the errors, the experimenter asked why he pointed at that animal, and Mike pointed at spikes and said "spikes". When he kept making errors, he did not want to continue the session. But, as mentioned, the component skills were still in place (tacting at the lower level, and producing the intraverbals). Therefore, Test 2 was administered despite Mike's poor performance in part of the maintenance session. It was anticipated that in Test 2, the instructional prompts may provide clues as to how he should sort the stimuli. It was thought that after production of the lower-level name (following the first sample-name prompt) it might be easier for him to produce the higher-level name (following the second prompt), because this situation was more like the intraverbal training, in which the child also produced the lower-level name first and then the higher-level name.

Category match-to-sample Test 2 with arbitrary stimuli. Before Test 2 with the arbitrary stimuli, Experimenter 1 conducted six review trials with the familiar stimuli, using the *tact-sample match-to-others* instruction. The child was prompted to tact the targeted familiar stimulus at both levels (e.g., reptile, animal), before making a selection. Mike reached the categorization criterion in the minimum of six trials. Then he was given the category match-to-sample Test 2 with the alien animals. The only difference with Test 1 was that in the test trials, as in the review trials, the *tact-sample match-to-others* instruction was employed. Now, the child was asked to tact the sample alien, at both levels (e.g. feb, and zaag), before selecting "the others".

After one trial block (of eight trials in total), Mike refused to play any further, and the session had to be discontinued. The testing during Block 1 progressed in the following way.

On Trial 1, when Mike was asked, "What's this?" he replied, "animal … lup" (it was indeed a lup). To the next prompt "And it's a … ?" he replied "lup-noom". But he was puzzled, when he was then asked for the others, and still puzzled when it was added, "it's a lup and a noom, can you give me the others?" Adding the more explicit question, "can you give me the other nooms?", made no difference. Therefore, Trial 2 was started.

The sample animal for Trial 2 (Hib 2) was held up, and Mike was asked what it was. He responded "hib". His reply to the second prompt "And it's a … ?", was "hib-zaag". However, when he was then asked "so is it a zaag or a noom?" he looked puzzled again and did not reply, but he did pick up the other hib (Hib 1).

In Trial 3, he tacted the targeted Tor 1, first as a tor, then as a noom. But when asked for the others, and later for the other nooms, he did not respond.

In Trial 4, he correctly tacted the sample stimulus as a feb. To the second prompt, he replied "feb-zaag". When asked for the others, and later for the other zaags, he did not respond, but he did pick up Feb 2.

In Trial 5, Hib 2 was targeted, which he tacted correctly at the lower level. To the second prompt he replied "hib-zaag", and he picked up Hib 1. He was then asked for the other zaags, and said "I really don't know" … following which he picked up the remaining zaags (Feb 1 and 2). So for the first time, he had selected the correct three aliens.

Finally, in Trial 6, the sample stimulus was Lup 2. In response to the first prompt, Mike tacted correctly ("lup"). To the second prompt, he correctly replied "noom". When asked for the others, he correctly picked up Tor 1, but no more (so not Lup 1 and Tor 2). To the question "Any more?", he replied "no".

Overall, in two trials in this block he selected no animals, and in two other trials he picked only the other lower-level category member. In one trial he correctly selected all three aliens (the other lower-level category member, as well as the other two higher-level category members). And in the remaining trial, he picked up only one animal, a higher-level category member.

After Block 1, Experimenter 1 gave the following instructions: "In this game it's all about names. There are quite a few zaags and quite a few nooms on the table. It's a secret whether you get it right. We'll tell you at the end." Mike seemed fine with this, but when Experimenter 2 then lined up the animals, Mike turned around on his chair, turning his back to Experimenter 2, and said he was tired. Friendly attempts were made to persuade him to continue, but he refused. And after this, he refused to play again.

Discussion

As in Study 1h, the aim of Study 2h was to investigate whether a child could sort the alien animal stimuli into categories solely on the basis of the common name (see one category member – select the others), without any direct categorization training. And if so, whether the child would show category sorting only at the lower name level, or also at the higher name level.

Although Mike did very well in pre-training with the familiar stimuli, in category match-to-sample Test 1 with the arbitrary stimuli (that is, the alien animals), he performed very poorly. Over the 48 trials he showed no correct category sorting at both the lower and the higher name level, and he made a correct selection at the lower level in only 29 out of 48 trials. As the vocalizations during the test indicate, Mike did produce lower-level names, or names very similar to them ("feb", and later "hippo" and "hippy", when referring to a hib), if only in a few of the trials. In the trials in which he tacted the sample animal at the lower name level before making a selection, he went on to pick up the other lower-level category member. In some trials it was not clear whether he was tacting the sample alien held up by the experimenter, or the alien(s) he selected. In one trial, he incorrectly picked up Hib 1 and Hib 2, and tacted them by saying "hippo". Interestingly, in response to the why-questions after the test, he produced lower-level names ("lups"), or names similar to them ("hippo's"), *and* he referred to physical features. The alien animals are arbitrary stimuli (i.e., stimuli that share no characteristics apart from the name; so they have no physical features such as shape or colour in common, that distinguish them from other stimuli). Therefore, Mike's explanations in terms of physical features, here and in the past few studies, are not valid. This is supported by the fact that given the same alien animal as sample on several trials he did not make consistent selections.

Horne and Lowe (1996) proposed that categorization may be brought about without direct training, by learning the same name to a set of arbitrary stimuli. In this case, there were lower-level and higher-level names for the stimuli. But the results, and especially the responses to the why-questions, suggest that Mike was not consistently basing his selections on either of these names.

Because he failed the category match-to-sample Test 1, Mike was then given the category match-to-sample Test 2 (also with the arbitrary stimuli), an altered version of the test, in which he was prompted to tact (see alien – say "name") the sample stimulus at both levels before selecting the other category members. Again, he did very well in pre-training, but not in the test with the alien animals. After six trials the test had to be terminated and Mike did not want to play again. In all six trials, when prompted, he tacted the alien correctly at the lower level. And when prompted further, he also tacted correctly at the higher level, in two trials. In the other four trials, he produced the correct intraverbal link (e.g., lup-noom) instead. However, the last prompt, "can you give me the others?" (and even "can you give me the other zaags/nooms?") seemed to really confuse Mike. He did not pick up any animals in two of the trials. In two other trials, he selected the other lower-level category member, and in another trial he also picked up just one alien, but this time it was one of the two higher-level category members. Finally, on one trial, Trial 5, he made the correct selection (the other lower-level category member, and the two other higher-level category members). It is puzzling that on Trial 2, in response to the prompt for the higher-level name, Mike responded by producing the intraverbal link (hib-zaag), but when he was then asked, "so is it a zaag or a noom?" he made no response. The prompts highlighted the necessity to base selections on names, thereby "breaking the strategy" Mike used so far, to go on physical features. That is, his behavioural repertoire of selecting the animals with physically similar features (or what he saw as physically similar features) was suddenly no longer appropriate.

When in a further instruction after the first six-trial block, he was further encouraged to use the names, the test became too aversive for him to continue.

Mike's ongoing focus on physical features of the animals ever since the transfer of function test in Study 2c, may have been a result of what is called a shape bias (see the general discussion in Chapter 5), or it may be connected to themes addressed in teaching sessions in the nursery. In the nursery, a different theme was chosen every so many weeks, and one of the themes had been animals (in general), while another theme had been jungle animals. In the weeks with these themes, the children were told stories and shown many pictures of animals/jungle animals, whereby there was a lot of attention for physical features of animals. For example, the children were shown that different kinds of giraffes had different patterns. Also, the features of different animals were described: how do animals differ? This may have had an effect on the behaviour of the children in the experimental sessions.

* * *

In the next chapter, the results of the first series of studies (1a-h) and the second series (2a-h) will be compared, and discussed further, also in relation to the literature summarized in the first two chapters of this thesis.

Chapter 4 – General Results and Discussion

This chapter will consider both series of studies described in the previous two chapters, in relation to each other, and in relation to the literature summarized in the first chapter of this thesis. Methodological matters and future directions will also be discussed.

Studies 1 and 2: a comparison

Tact training: Studies 1a and 2a. In Studies 1a and 2a, the children received vocal tact training with arbitrary stimuli (the alien animals), which was followed by testing for listener behaviour. In Study 2a, pre-training with familiar stimuli preceded this. Figure 4.1 shows the average number of trials for pre-training in Study 2a, and for arbitrary stimulus training both in Studies 1a and 2a, along with the average number of trials per stage within the alien tact training procedure.

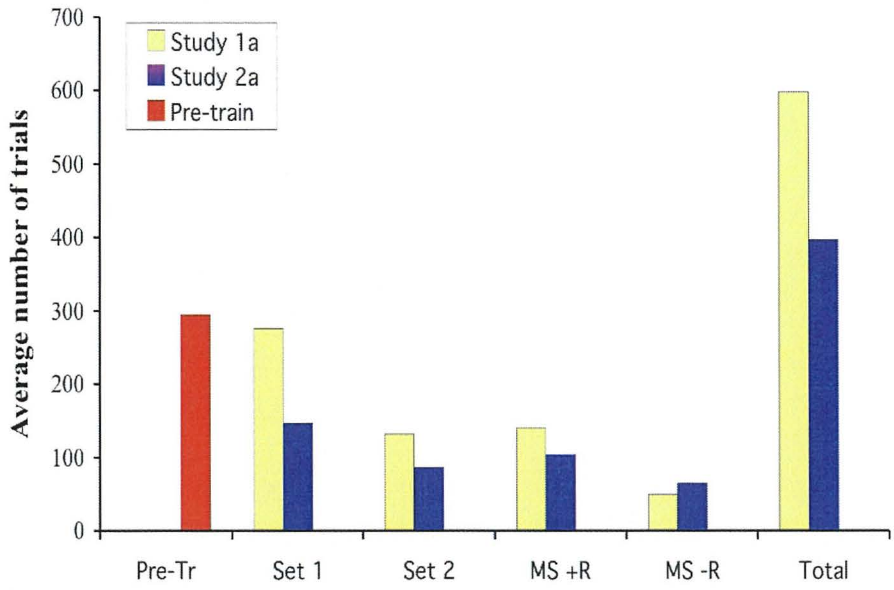


Figure 4.1 The average number of trials for pre-training (in red) with familiar stimuli in Study 2a, along with the average number of trials per stage in alien animal tact training for children in Study 1a who had no pre-training (in yellow), and for children in Study 2a who did receive pre-training (in blue). Also depicted is the average number of trials the children needed in total in tact training with the aliens, in both studies.

As can be seen, the children who completed alien tact training in Study 1a needed on average 598 trials (range 372—984), whereas the children in Study 2a who had received pre-training required only 397 trials (range 252—600).

In Study 1a the average number of trials needed to learn to tact Set 1 was 276 (range 168—396), but only half that number (132: range 36—432) were required for Set 2. Similarly, in Study 2a the average number of trials for Set 1 was 147 (range 60—228), and for Set 2 it was 86 (range 36—252). For the mixed sets with reinforcement phase of training an average of 140 (range 96—228) trials were required in Study 1a, compared to 104 (range 96—132) in Study 2a. Under no reinforcement conditions for the mixed sets, 50 (range 24—72) trials were needed in Study 1a, compared to 65 (range 48—192) in Study 2a.

Both studies show clear evidence of a learning set effect (Harlow, 1959)[1] for tact training. Comparison across studies suggests that pre-training with familiar stimuli decreases by half the number of tact training trials required with the alien stimuli. Likewise, within each study the children learned to tact the Set 2 alien stimuli much more quickly than they did Set 1. In the mixed sets phases, possibly because performances approached the minimum number of trials, the advantage of pre-training was no longer evident in comparisons of Study 1 with Study 2. This overall outcome indicates that, in general, experience with the procedures, including the instructions plays a role in how rapidly children learn, even in seemingly simple learning situations (see Horne et al., 2006, p.268).

The total number of trials to complete training for females was on average 615 (range 528—900) in Study 1a, and 510 (range 384—600) in Study 2a. Males, on average, required 588 (range 372—984) trials in Study 1a, compared to 347 (range 252—432) in Study 2a. One possible reason for this sex difference will be considered later, in the Discussion section of this chapter.

Data for individual children in both studies are shown in Figure 4.2. When, as here, the number of trials is presented as a function of age, it is clear that older children do not learn in less trials than younger ones in either of the studies.

In Study 1a, where no pre-training was given, the child who learned in the least trials (372) was Jim. Of the 13 participants in Study 2a, in which pre-training was given, 7 needed less trials-to-criterion than Jim. Furthermore, in Study 2a, the child who required the most trials (600) was Ginny. In Study 1a, 7 children needed more trials than Ginny (and even then, 2 of them did not reach criterion).

---

[1] A learning set effect is the effect that learning takes place progressively more quickly, when similar learning tasks are encountered in similar situations. For example, problem solving tasks or

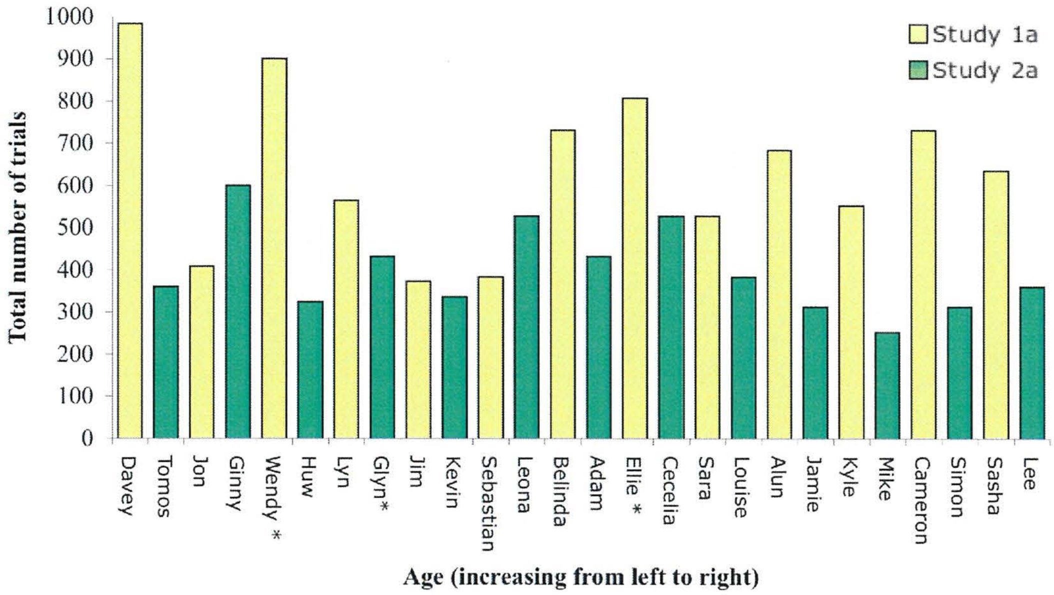


Figure 4.2 The total number of trials needed to reach criterion over all stages in alien tact training until the listener behaviour test, for all children in Study 2a (in green), with the youngest child on the left (Tomos, 3/3) and the oldest child on the right (Lee, 4/4). For comparison, the data for the children in Study 1a are shown as well (in yellow). (*Child did not complete the study.)

Figure 4.3 and Figure 4.4 show the number of trials for female and male participants separately, in both studies. Overall, both females and males in Study 2a needed fewer trials than in Study 1a. As a group, in both studies the males needed fewer trials to criterion than the females. Within Study 2a, only 1 girl – Louise – "scored" within the range of the boys (she had 384 trials); only 2 boys needed more trials than she had. For both females and males, the ranges for the total numbers of trials are much smaller in Study 2a than in 1a. In Study 1a, the difference between the girl with the fastest and the girl with the slowest learning process was roughly 400, while for boys this difference was about 600. In comparison, in Study 2a this difference was roughly 200 trials for both sexes.

---

discrimination tasks may take considerable time to complete when first encountered, but much less

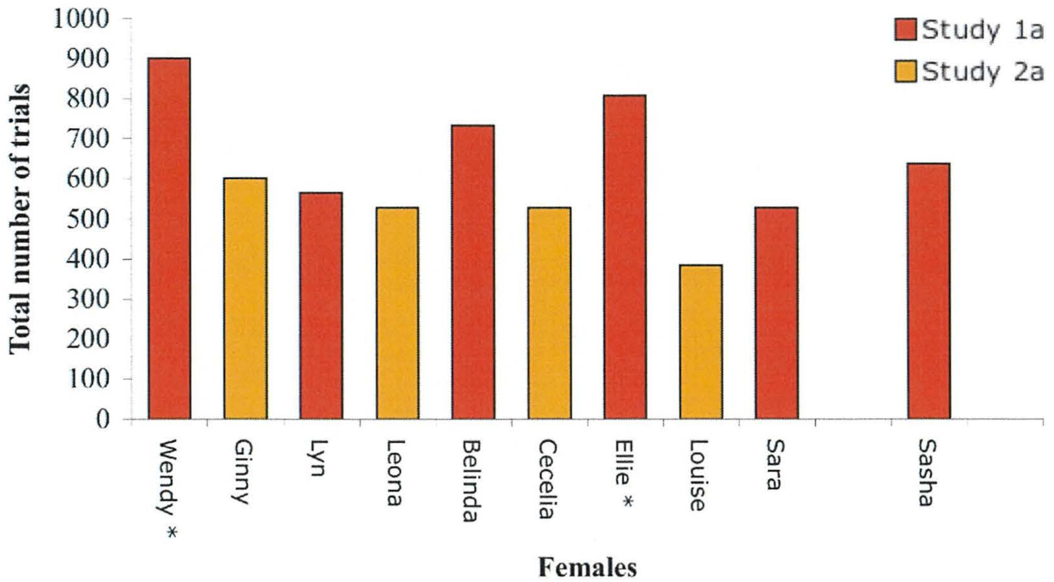


Figure 4.3 The total number of trials for each female in Study 2a (in orange), and in Study 1a (in red). (*In Study 2a there were 2 female participants less than in Study 1a.)

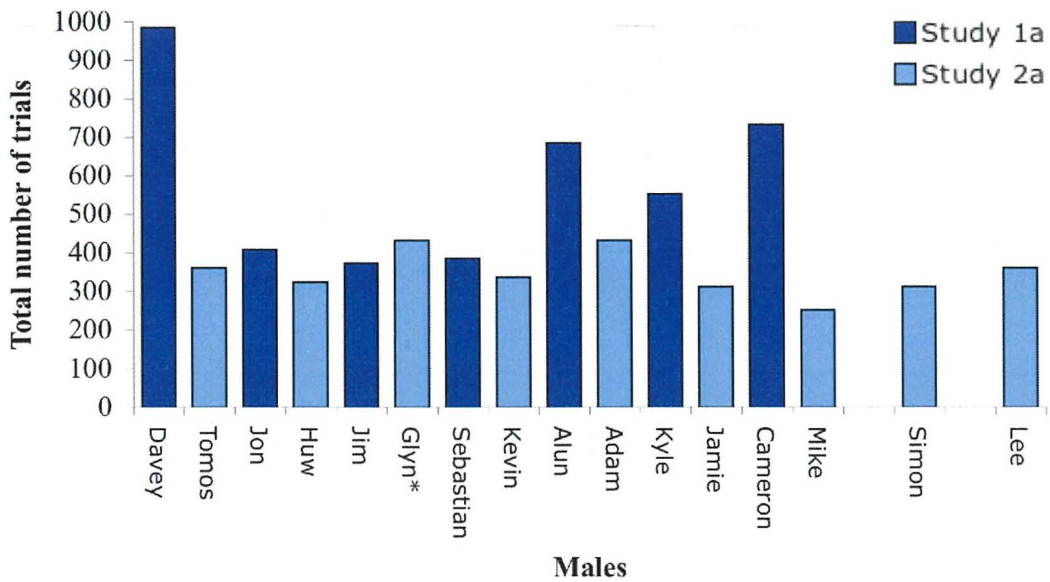


Figure 4.4 The total number of trials for each male in Study 2a (in light blue), and in Study 1a (in darker blue). (*Child did not complete training. NB. In Study 2a there were 2 more male participants than in Study 1a.)

---

time once a few of these tasks have been completed (see also Schrier & Thompson, 1980).

Listener testing: Study 1a and Study 2a. After alien tact training the children in both studies were tested for the corresponding listener behaviour. In Study 1a, of the 13 children who participated, 11 completed tact training, and all 11 passed the listener behaviour test. Of the children who passed, four showed an errorless performance over all 48 trials in the test (24 trials per mixed set).

In comparison, of the 13 children who participated in Study 2a, 12 completed training and were given the listener test, which all passed. Six children showed an errorless performance, while 2 children made 1 error, and 3 children made 2 errors.

Unlike the tact training data, there were no differences across studies in performance on the listener test.

Intraverbal training: Study 1b and 2b. In Study 1b and Study 2b an echoic and intraverbal word game was introduced, linking each of the lower-level names hib, feb, tor, and lup with a potential higher-level name, either zaag or noom. Following this, the children were tested for listener behaviour with regard to the potential higher-level names. In Study 2b, pre-training with the names of the familiar stimuli preceded this procedure. Figure 4.5 shows the number of intraverbal training trials with alien names, in both Study 1b and Study 2b, per child. It can be seen that for Leg 1 of intraverbal training all children in Study 1b needed approximately 100 training trials. In Study 2b, only 1 child (Adam) had less than 150 trials, whereas 2 others (Tomos, and Ginny) needed between 150 and 200 trials. Two more children (Huw, and Lee) needed at least twice the number of trials of the children in Study 1b, while 2 other children (Louise, and Mike) needed three times that number, and 1 child (Simon) even needed six times that number of trials.

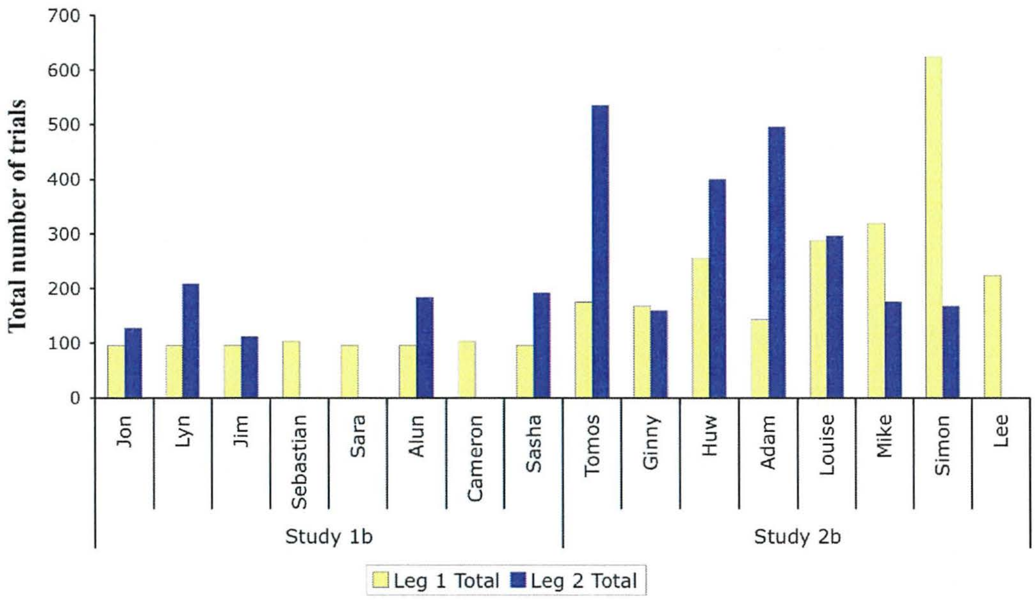


Figure 4.5 The total number of training trials for each participant in the intraverbal word game (Leg 1 in yellow, Leg 2 in blue). The participants in Study 1b, who did *not* have pre-training with the familiar name links, are shown on the left (Jon - Sasha), whereas the participants in Study 2b, who did receive pre-training, are shown on the right (Tomos - Lee). Sebastian, Sara, Cameron, and Lee were not available for training and testing in Leg 2.

For Leg 2 in Study 1b, all children needed more trials than for Leg 1. For 2 children (Jon, and Jim) the difference was not very big, but the other 3 children needed approximately twice as many trials as for Leg 1. In Leg 2, Study 2b, 3 children (Ginny, Mike, and Simon) needed between 150 and 200 trials, which is within the range of the children in Study 1b. Of the other 4 children in Study 2b, Louise needed about 100 trials more than the slowest children in Study 1b, whereas Huw needed 200 trials more than Louise, and Tomos, and Adam required around 100 trials more than that.

So overall, the children in Study 2b were slower and in most cases much slower, than the children in Study 1b in learning the intraverbal relations between the lower-level alien names at the potential higher-level names. There are several reasons, which could account for this difference. It should be noted that the added pre-training for Study 2a, and especially Study 2b, extended the experimental procedure considerably, for most children. Although the children enjoyed the experimental sessions throughout, particularly the one-to-one interaction, over time the experimental part of the sessions became less reinforcing, mostly because of repetition of the procedure (in some cases there seemed to be no end to the word game trials; e.g., for Adam, Tomos, and Simon, who had 500 trials or more of intraverbal training for one leg). This repetition made children less compliant, which in turn had a negative effect on their performance on the trials. Another issue is that it seemed that the more intraverbal relations the children learned, the more they mixed them up. The children in Study 1b learned four intraverbal relations, whereas the children in Study 2b learned eight: four in pre-training, and four in alien name training. Sometimes children in Study 2b also mixed up the intraverbals in alien name training with those in pre-training, saying for example, "feb-animal-feb-animal", and, during maintenance training for the intraverbal relations with the familiar names, "fish-noom-fish-noom".

Higher-level listener behaviour test. After intraverbal training with the alien names, the children were tested for listener behaviour at the higher level, first after Leg 1 intraverbal training and then again after Leg 2. For Leg 1, in Study 1b 5 out of 8 children passed the test, whereas for Study 2b, 7 out of 8, and the remaining child passed when retested (in several sessions). For Leg 2, 3 out of 5 children passed in Study 1b, but for Study 2b, 3 out of 7 children passed. So for Leg 1, the children in Study 2b (who had pre-training) did better on the test, as a group, than the children in Study 1b (who did *not* have pre-training), but for Leg 2, although the same number of children passed the test in both studies, the children in Study 1b did relatively better, as a group.

In both studies together, for 13 (out of 16) children it was enough to learn the intraverbal relation between the lower-level names and potential higher-level names of Leg 1, to also acquire listener behaviour at the higher name level, without direct training. For 6 (out of 12) of these children, this also applied to Leg 2.

It is unclear why children would pass the listener behaviour test for one leg and fail it for the other leg. In both studies, extra sessions in which the children who failed the test were given more explicit rules (e.g., a hib *is a* zaag) in a further word game did not help them to pass retests. Because the component skills of tacting at the lower name level, and producing the intraverbal relations, were still in place, it seemed that the children simply did not respond in terms of the intraverbal relations, or the explicit rules stating the higher order category relation, during the listener behaviour test. Also, in both studies consistent reversal of responses was found for some children. In the post-test interviews in this study and later studies, many children indicated that their responses were based on physical features of the aliens. So although the alien animals are arbitrary stimuli such that few of them have any physical features in common, the children appeared to try to base their listener responses on physical features, as they learn to do for most real-life animals.

For the children who failed the listener behaviour test at the higher level, even after the extra sessions with more explicit rules, it would presumably be necessary to explicitly tell them the higher-level names in the presence of each and every one of the alien animals.

Novel gesture training: Study 1c and 2c. In Studies 1c and 2c, two responses (gestures, comparable to greetings) were taught, one to a member of one higher-level category (a zaag), and one to a member of the other higher-level category (a noom). Then the children were tested for name-based transfer of these responses to the aliens that were not involved in training. In Study 2c, pre-training with familiar stimuli preceded this procedure.

The number of trials needed to train the novel gestures with the alien animals was roughly the same in both studies; the children in Study 1c had on average 64 trials (range 40-88), while the children in Study 2c had 56 (range 40-88).

Transfer of gestures test. The transfer of function test in both studies gave the same results. In both studies, 1 out of 3 participants showed full transfer of function (at the lower, as well as at the higher level), 1 child showed partial transfer of function, to one pair of aliens, and the third child showed transfer to none of the pairs. In the two cases of partial transfer (one case in each study), transfer concerned one alien belonging to the same lower-level category, and one alien belonging to the same higher-level category as one of the aliens used in training. With the less stringent criterion, at the level of the individual aliens, slightly more partial transfer was found, which is also the same in both studies. Therefore, the added pre-training with familiar shapes, in Study 2c, did not help the children to perform better on the alien transfer of function test; the transfer found is not more consistent for either level.

Novel vocal behaviour training: Study 1d and 2d. In Studies 1d and 2d, two new responses (animal cries) were trained, one to a zaag, and one to noom. After training, the children were tested again for name-based transfer of these responses to the aliens that were not involved in training. There was no pre-training with familiar stimuli in Study 2d. Therefore, Study 2d was an exact replication of Study 1d.

In vocal behaviour training, all children in both studies (that is, 1 child in Study 1d, and 3 in Study 2d) reached criterion in the minimum of 40 trials.

Transfer of vocal behaviour test. The transfer of function test in both studies gave similar results again, in the sense that none of the children in either study excelled. Rather, there was deterioration in the children's performance, compared to that in the transfer of function test for gestures in Study 1c and Study 2c. None of the participants in Study 1d and Study 2d showed full transfer of function at the lower and the higher level. The child in Study 1d showed partial transfer of function, to one pair of aliens, whereas the 3 children in Study 2d showed transfer to none of the pairs, while one of them even showed complete reversal of responses. The partial transfer shown by the child in Study 1d concerned one alien belonging to the same lower-level category, and one alien belonging to the same higher-level category as one of the aliens used in training. With the less stringent criterion, at the level of the individual aliens, Simon (in Study 2d) did seem to show transfer to one alien.

In retests, some children did better; Jon in Study 1d, reached criterion for two pairs of aliens (two lower-level and two higher-level category members), while in Study 2d Huw showed transfer to one pair, and Simon still to none of the pairs. With the less stringent criterion at the level of the individual aliens, Huw shows transfer to four aliens and Simon to two, which for Simon was completely, and for Huw partly, due to his repeated production of "Raagh!" responses. Mike, in Study 2d, was not retested, because his responses were completely consistent (he showed complete reversal of responses).

So in the test, as well as the retest, Jon in Study 1d performed better than the 3 children in Study 2d. And just as the added pre-training with familiar shapes in Study 2c did not help the children to perform better on the alien transfer of function test, previous experience with function training and transfer of function testing (in Study 1c and 2c, respectively) did not make for better performance in the tests of Study 1d and 2d; deterioration in performance was found rather than improvement.

All cases of partial transfer at the higher level that were found in both studies could be due to transfer at the lower name level, in combination with "selection by exclusion", because all the aliens for which transfer was found at the higher level had been presented with aliens belonging to the same lower-level category as the aliens employed in function training.

Transfer of function test with generic instruction: Study 1e and 2e. Studies 1e and 2e were tests only, focussing on the effects of different verbal prompts on the child's response transfer. The verbal prompt employed, "What *can* this one *do*?" was more general than the prompts in the previous two studies (i.e., "How does this one go?" and "What does this one say?"). Would this more general verbal prompt have stimulus control over both types of previously taught functional responses (gestures and cries) in the absence of direct training? Because there was no pre-training in Study 2e, it was an exact replication of Study 1e. At this stage, in each of the series of studies there was only one participant left (Jon and Mike, respectively).

In Study 1e, the more general verbal prompt did indeed evoke both types of responses, but never in the same trial. That is, in the first three trial blocks Jon produced only gestures to the presented alien pairs, thereby showing perfect transfer of function at the lower and the higher level. In the other trial blocks (administered in different sessions) he produced gestures to some aliens, and animal cries to the others. He produced gestures to the aliens that had been involved in gesture training, as well as to the aliens in the same lower-level category, and he produced cries to the aliens that had been involved in training for the cries, as well as to the the aliens in the same lower-level category as these. All Jon's responses were correct.

In Study 2e, the more general verbal prompt did not evoke gestures, only animal cries. And with regard to these, Mike reached criterion for only one of the four pairs of aliens, and one of the aliens in this pair had been part of function training for the animal cries. For the other aliens, Mike showed reversal of responses for five of the remaining aliens, and a tendency to do the same for the last alien animal. One of the aliens he showed reversal for, was the other alien that had been involved in training for the animal cries.

Listener behaviour to gestures test: Studies 1f and 2f. Study 1f tested for listener behaviour in response to models of the gestures trained in Study 1c. The listener test was limited to the two aliens that had been involved in training of the production of the gestures. Study 2f replicated this. No pre-training was added in Study 2f, but the testing was extended to all aliens. Therefore, the two studies can only be compared in part.

The studies had similar results: both Jon (in Study 1f) and Mike (in Study 2f) performed well; Jon had 90% and 100% correct, while Mike had 100% correct for each of the two aliens. So once the children had learned to produce the gestures to these aliens, listener behaviour was shown to be in place.

Listener behaviour to trained vocalisations test: Studies 1g and 2g. Study 1g tested for listener behaviour with regard to the animal cries trained in Study 1d. The listener test was limited to the two aliens that had been involved in training of the production of the animal cries. Study 2g was a replication of this. No pre-training was added, but (as in Study 2f) the testing was extended to all aliens. Hence, only partial comparison of the two studies is possible.

Similar results were found again: Jon (in Study 1g) had 100% correct for each of the two aliens, while Mike (in Study 2g) had 90% correct for each. Therefore, as was the case with the gestures in Studies 1f and 2f, once the children had learned to produce the animal cries to these aliens, listener behaviour was shown to be in place.

Category match-to-sample tests: Studies 1h and 2h. Studies 1h and 2h tested whether the children would now be able to categorize the aliens on the basis of common names, without direct training. Would the children show category sorting at the lower and the higher level? A category match-to-sample test was administered to investigate this. In Study 2h, pre-training with familiar stimuli was added in which the child was trained to respond correctly to the instructions later used in category testing with the aliens, and to sort the familiar stimuli into categories.

In Study 1h, Jon performed very well, showing correct category sorting at the lower as well as the higher name level. He passed the category match-to-sample test, having 100% correct responses for three of the aliens, and 83% (5 out of 6 selections) correct for each of the other five aliens. When he made an incorrect selection, on five out of 48 trials, he did always pick the correct lower-level category member.

In Study 2h, Mike performed well in pre-training with the familiar shapes, reaching criterion in the minimum number of trials in two of three phases. But in the category match-to-sample test with the aliens he did not produce any correct category sorts. When he picked the correct three aliens, on one trial, he selected an additional incorrect one. And he was not consistent in the aliens, or even the number of aliens, he selected for each of the sample stimuli. However, he did pass for selection of the other lower-level category member; he picked the correct lower-level category member in 29 of the 48 trials (while the criterion was 16/48, $p < .001$).

Because Mike failed the test, he was given an alternative version of the test, which required him first to tact the sample stimulus at both levels, and then to select the other category members. Here as well, he was given pre-training first, to ensure correct responding to the *tact-sample match-to-others* instruction. He reached criterion in the minimum number of trials. However, when tested with the aliens, he refused to play the game any further after six trials. In these six trials, he tacted the target stimulus correctly at the lower level. In two of these trials, he also tacted correctly at the higher level, while in the other four he produced the correct intraverbal link. But he made the correct selections in only one trial. In two trials he picked up the correct lower-level category member, and no other aliens.

In sum, the added pre-training did not result in better performance. On the contrary. There was a striking difference between the 2 participants in Study 1h and 2h, with Jon, who had *not* had pre-training, showing a near perfect performance, whereas Mike, who had been given pre-training, performed considerably below the criterion level that was set for the test. And tacting the sample stimulus before selecting any aliens did not improve Mike's categorizing behaviour.

Production of gestures and animal cries, listener behaviour, and transfer of function: a comparison

Studies 2c and 2f. In Study 2c, the children were taught to produce novel responses (gestures) to two aliens, and then they were tested for name-based transfer of these responses to the aliens that were not involved in training. In training, as well as in testing, they were required to *produce* the gestures. Study 2f, was a test for *listener behaviour* with regard to these gestures; Experimenter 2 produced the gestures, and asked the child to point to the alien that "goes like this". In effect, Study 2f tested for (untrained) listener behaviour for the aliens involved in gesture production training, and it tested for transfer of this untrained listener behaviour to the other aliens. So Study 2c tested for transfer of the production of the gestures, while Study 2f tested for transfer of listener behaviour with regard to the gestures. Figure 4.6 shows these transfer of function data together.

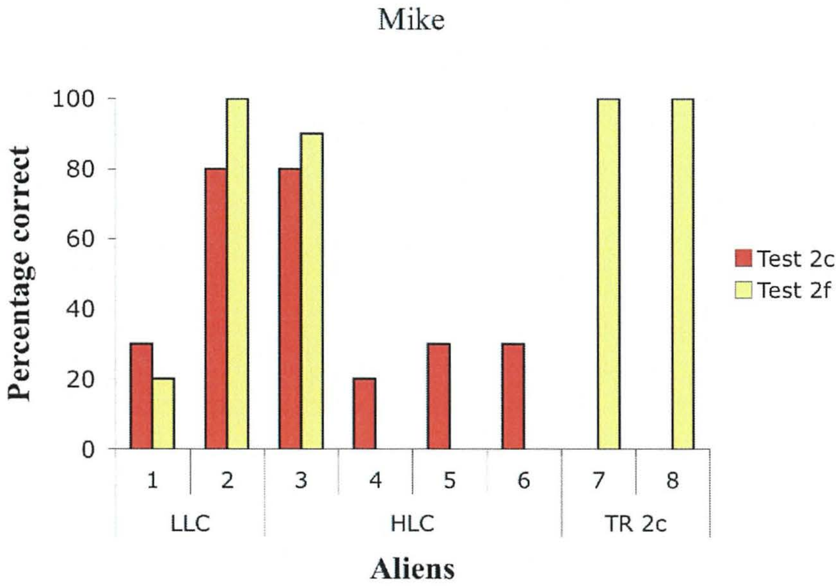


Figure 4.6 The percentage of correct responses for all eight aliens in the test for listener behaviour and transfer of function in Study 2f (in yellow), and for the six aliens in the test for transfer of function in Study 2c (in red). Aliens 7 and 8 had been used in function training in Study 2c. Aliens 1 and 2 were of the same lower level category (LLC), while Aliens 3-6 belonged to the same higher level category (HLC) as the aliens used in this training. The additional transfer of function criterion was 80% (8 out of 10) correct for each alien.

In the transfer of gestures test of Study 2c, six rather than eight aliens were used. The other two had been used in function training. For the six aliens used in both tests, the picture is one of reasonably consistent responses. Mike reached criterion for the same two aliens (2 and 3) at the individual alien level, and reversed responses where he did so before, or where he showed a tendency to do so earlier.

Studies 2d and 2g. In the transfer of vocalisations test of Study 2d, the children had learned to produce another set of novel responses (animal cries) to two aliens, and following that they were tested again for name-based transfer of these responses to the aliens that were not involved in training. As with the gestures, in training, as well as in testing, they were required to *produce* the animal cries. Study 2g, was a test for *listener behaviour* with regard to the animal cries; Experimenter 2 produced the cries, and the child pointed to the alien that "said that". So, Study 2g tested for (untrained) listener behaviour for the aliens involved in animal cry (production) training, and it tested for transfer of this untrained listener behaviour to the other aliens. Therefore, Study 2d tested for transfer of the production of the cries, while Study 2g tested for transfer of listener behaviour with regard to the cries. Figure 4.7 shows the transfer of function data together. In the test of Study 2d, six aliens were used. The other two had been part of function training. Mike's responses for all six aliens were completely consistent within and across the two studies: full reversal was found.

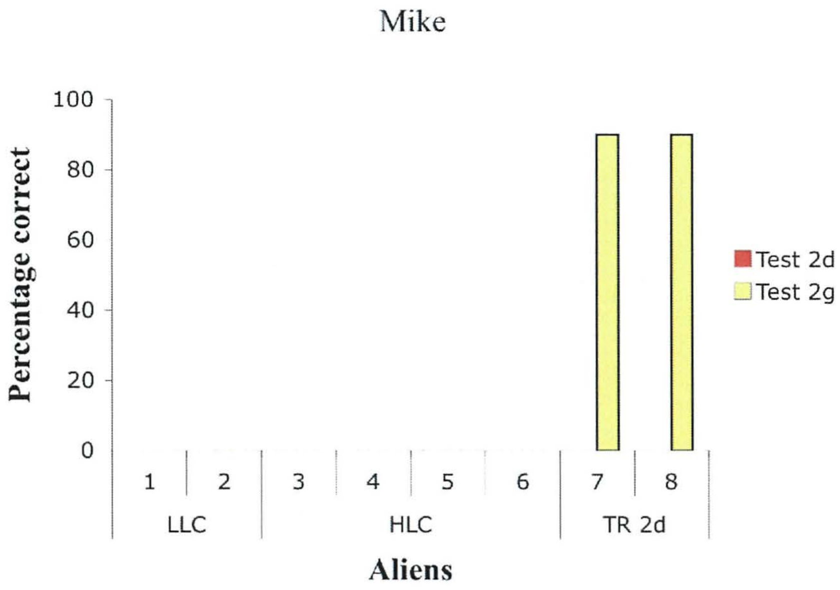


Figure 4.7 The percentage of correct responses for all eight aliens in the test for listener behaviour and transfer of function in Study 2g (in yellow), and for the six aliens in the test for transfer of function in Study 2d (in red). Aliens 7 and 8 had been used in function training in Study 2d. Aliens 1 and 2 were of the same lower level category (LLC), while Aliens 3-6 belonged to the same higher level category (HLC) as the aliens used in this training. The additional transfer of function criterion was 80% (8 out of 10) correct for each alien.

The experimental data and the literature

As cited in Chapter 1, Horne and Lowe define naming as "a higher order bidirectional behavioural relation that (a) combines conventional speaker and listener behavior within the individual, (b) does not require reinforcement of both speaker and listener behavior for each new name to be established, and (c) relates to classes of objects and events" (1996, p. 207). Horne and Lowe explain with regard to Part b of this definition that no *separate* reinforcement is required for speaker and listener behaviour; in the course of speaker behaviour training (tact training) a child's listener behaviour will indirectly be reinforced as well[2]. They also point out that for a tact to emerge during listener training, the child must be able to echo the listener stimulus. Finally, the authors describe the implications of this definition for determining whether children can name.

One way to test whether children can name, related to Parts a and b of the definition, is to train speaker behaviour (tacting) in relation to a novel stimulus, and then to test for the corresponding listener behaviour. Both studies described in Chapter 3 and Chapter 4 of this thesis, started with the children learning to tact the newly designed aliens. When tested for listener behaviour following this, all 23 children who were tested passed. One can therefore say, that all children had learned to name the alien animal stimuli. This is in line with findings from previous studies (see Lowe, et al. 2002; and Lowe, et al. 2005). Together, these studies provide

---

[2] Tact training for a particular object requires the child to look at the object when its name is first produced by a parent or experimenter, and later when hearing the verbal prompt, "What's this?" The child then says a name, and receives feedback. In this process, seeing the object precedes the child saying the name, and often the child will still look at the object when saying the name and receiving feedback (or the child may look at the stimulus again, while she is reinforced for tacting it). In that situation, reinforcement for tacting is provided while the child is showing listener behaviour (looking at the object), which will strengthen the listener behaviour.

support for the naming account, which suggests that when a child is trained to tact an object, in the course of this the child is likely to also learn the corresponding listener behaviour, without it being separately trained (see Horne & Lowe, 1996, p. 201; and see part b of the definition of naming above). This would apply to normally developing children of 2- to 4-years old (see Lowe, et al. 2002). The naming account proposes that listener behaviour may be learned without separate training because by the time the child has sufficient echoic behaviour to be able to learn a particular tact response she will have already learned a considerable listener repertoire and so upon hearing the trainer's object label she will orient to the referent object as she echoes the label: the praise that follows her echoic response will serve to reinforce her listener response as well as her echoic response in the training interlude.

One child, Davey reached the overall mastery criterion, but failed to meet the individual mastery criterion for one of the eight aliens. Interestingly, he also had considerable problems learning to tact this particular alien, as explained in Chapter 3. So there was variability in both his speaker and his listener behaviour with regard to this stimulus.

In their description of the development of intraverbal behaviour, Horne and Lowe point out, that "when names are reliably linked within an intraverbal sequence, listener behavior of any one of the name relations will be increasingly evoked by the other name relations and vice versa" (1996, p. 210).

Support for this was found in Studies 1b and 2b. After having learned to tact the alien animals at the lower name level in Studies 1a and 2a, and having demonstrated appropriate listener behaviour at the lower level in the tests that followed, 16 children were taught a word game that trained an intraverbal link between the lower-level names (hib, feb, tor, and lup) and potential higher-level names (zaag, and noom). Once these responses were reliably linked within an intraverbal sequence, the children were tested for transfer of listener behaviour to the potential higher-level names, that is, for selection of the correct animals when only presented with the potential higher-level names. In the test for Leg 1 of the hierarchy, 13 of the 16 children showed correct selection of the animals, which could be interpreted as transfer of listener behaviour through intraverbally linked names, as proposed by Horne and Lowe. In the test for Leg 2, this was found for 6 of the 10 children.

In addition to there being transfer of listener behaviour from one name level to another, one can also speak here of an "intraverbal means of (...) establishing functional equivalence between physically different stimuli" (Horne & Lowe, 1996, p. 210); for each child the name "zaag" was linked with two lower-level names, for example, hib and feb, and therefore the aliens with the names hib and feb become functionally equivalent, both in a test for listener behaviour at the higher name level, and in category match-to-sample tests (Study 1h, and Study 2h) where the child is shown, for example, one zaag, and asked for the others. In Studies 1a and 2a functional equivalence was established between physically different stimuli as well. Once the children had learned that in Set 1 one alien was a hib, one a feb, another one a tor, and the last one a lup, they were presented with a new set, also consisting of a hib, a feb, a tor and a lup. The two hibs look different from each other, and so do the two febs, tors, and lups. By teaching the children the same name for these physically different stimuli, the stimuli become functionally equivalent. When a child learns names at a higher level (like "furniture", for chairs), Horne and Lowe (1996) speak of "a more general level of functional equivalence", and they add, "in this way, naming, incorporating as it does bidirectional relations between a single verbal response and a class of objects or events, can establish and maintain more than one level of functional equivalence" (1996, p. 205). As noted, in Studies 1b and 2b the higher level of functional equivalence was established in some children by means of intraverbal training.

For the children who did not pass the test for listener behaviour at the higher name level, review trials before the test, and also review trials in a session after the test usually showed that the component skills (tacting at the lower level, and the intraverbals, and after additional sessions also the more explicit rules like, for example, "a hib is a zaag") were still in place. So it seemed that the behaviour of these children during the listener behaviour test at the higher name level was not under the control of the intraverbal relations they had learned, which was supported by Sara's reaction to one of the clues given during the test to remind her of the intraverbals: "Is that important??" This suggests that it may be necessary for these children to be prompted to produce each of these component "skills" on each trial of the test for listener behaviour at the higher level, by first prompting the child to tact one of the presented stimuli at the lower level, then to prompt for the relevant intraverbal, to repeat this for the other stimulus, and finally to prompt for listener behaviour at the higher name level by asking, "so which one is the zaag/noom?" With Lyn, in Study 1b, something similar to this (prompting for lower-level tact, then for intraverbal, and finally for listener behaviour at the higher level) was done in additional sessions, when she was given feedback, with little effect on her performance. Perhaps the prompts for the component skills should be given before prompting for higher-level listener behaviour, as described above, to have any effect. If that still does not establish listener behaviour at the higher level, then it would presumably be necessary to explicitly tell the children the higher-level names in the presence of the alien animals.

One defining characteristic of naming is that it relates to *classes* of objects and events (Horne & Lowe, 1996, p. 207 – Part c of the definition). And Horne and Lowe (1996) proposed that stimulus classes may be established by the child simply learning a common name for several stimuli. Lowe et al. (2002, and 2005), and Horne et al. (2004) found support for this, and concluded that common naming is "highly effective" in bringing about stimulus classes. The stimuli that form a class may look very similar (e.g., cats), or vary quite considerably though still showing physical similarities (e.g., dogs), or show no physical similarities that distinguish them from stimuli in other classes (e.g., furniture). When stimuli that form a class show no physical similarities, we speak of arbitrary stimulus classes. It is especially these arbitrary stimulus classes that have been the focus of behaviour analytic research, and it is for the establishment of these arbitrary stimulus classes that naming may be critical (see Horne & Lowe, 1996; and Lowe et al., 2002, and 2005).

The alien animal stimuli employed in the present studies are arbitrary stimuli; the aliens belonging to one class (e.g., the hibs, or at a higher name level, the nooms) have no physical characteristics in common, like shape or colour, that set them apart from the stimuli in other classes (e.g., the febs, or at a higher name level, the zaags). Therefore, the only thing they have in common is the name.[3]

In line with Horne and Lowe (1996), Lowe et al. (2005) listed two ways of investigating the functional properties of common name or category relations: (1) by

---

[3] Note that for each child, the names were randomly assigned to the animals, so if the two hibs for one child happened to both have spikes, that was a coincidence, and that would not be the case for other children.

testing for transfer of function within but not between members of common-name relations, and (2) by testing category sorting among sets of arbitrary stimuli with common names. Both transfer of function and category sorting were tested in the present studies.

Name-based transfer of function from one member of a stimulus class to the other class members (e.g., from one hib to the other, or from one zaag to another), but not to members of other classes (e.g., not from a zaag to a noom) was the focus of Study 1c, Study 2c, Study 1d, and Study 2d. Before participating in these studies, the children had learned to tact the aliens at the lower name level (in Study 1a and Study 2a), and had learned intraverbal behaviour linking the lower-level names with the potential higher-level names (in Study 1b and Study 2b). In addition, they had demonstrated appropriate listener behaviour at the lower and the higher level in subsequent tests.

In Studies 1c and 2c, the children were trained to produce two different gestures, one to a zaag, and one to a noom. This was followed by a test for transfer of function to the other class members, at the lower and higher level. Of the 6 children who participated, 2 (Jon and Simon) showed full transfer of function to the other lower-level category member, and the other higher-level category members. Two children showed partial transfer of function, that is, to one pair of aliens, and 2 children showed transfer to none of the pairs. The partial transfer concerned one alien of the same lower-level category, and one alien of the same higher-level category as one of the aliens used in training. Of the 2 children who showed full transfer, only Jon (in Study 1c) could explain his behaviour when asked, and he did so by referring to the alien names, but only to the lower-level names. In Study 1c, Alun showed transfer to one alien pair, but his vocalisations during the test suggest that he based his responses on shape (comparing the shape of the gestures with the shapes of the aliens). Sasha showed transfer to none of the alien pairs. Interestingly, when on one trial she produced the intraverbal (consisting of both the lower and higher-level name) for that animal she also produced the correct gesture. After the test, Sasha was given additional function training in which she was required to tact the aliens at both levels before producing the gesture. Following this, she was tested again, and in the test she was also prompted to tact the aliens at both levels before producing the gesture. With this procedure, Sasha did much better; she now showed transfer to two alien pairs. Unfortunately, time constraints did not allow for implementation of this procedure with the other children in Studies 1c and 2c who showed partial or no transfer of function. There were no vocalisations during the test in Study 2c. In training for this study, Mike did come up with a name for the "hands in front" gesture; he called it "bridge". He did not utter this name overtly during testing.

Studies 1d and 2d were further investigations into transfer of function. This time the children were trained to produce two different animal cries, one to a zaag, and one to a noom. Then they were tested again for transfer of function to the other class members, at the lower and higher name level. None of the 4 children who participated, showed full transfer of function at the lower and the higher level. One child (Jon, in Study 1d) showed transfer of function to one alien pair (one alien of the same lower-level category, and one of the same higher-level category as the aliens involved in training), while the other 3 children (Huw, Mike, and Simon) showed transfer to none of the pairs. Mike even showed full reversal of responses. Three children (including Jon) were retested, and two of them showed improved performance; Jon now reached criterion for two pairs of aliens (two lower-level and two higher-level category members), while Huw showed transfer to one pair. Jon's vocalisations during the test suggest that his behaviour was based on names. Several times during the test he produced lower-level names, and once an intraverbal. Twice he indicated that his production of a animal cry depended on lower-level names and selection by exclusion (when Lup 1 and Hib 2 were presented and the lup was targeted, he said, "Raagh, because I know the hib says Boo!" and when Feb 1 and Tor 2 were presented and the tor was targeted, he said, "Boo! Because that one (Feb 1) says Raagh!"). There were not many vocalisations during the test in Study 2d, and none that give any clues about the use of names by the children.

Studies by Lowe et al. (2005) and Horne et al. (2006) investigated differences in transfer of function after tact training and after listener behaviour training, and found that children showed transfer of function after tact training, but not after listener behaviour training (unless children passed a tact test showing that tacting was also learned in the course of this listener behaviour training, without it being directly trained). It should be noted that no different name levels were involved in these studies.

As pointed out in Chapters 3 and 4, the children in Study 1c, Study 2c, Study 1d, and Study 2d can name at the lower level, but we cannot be certain that they can also name at the higher level. They had intraverbal training linking lower-level names and potential higher-level names, and were tested for listener behaviour at the higher level. All children who participated in the c and d studies passed this listener behaviour test at the higher level.

One question that arises is whether they passed the listener test by responding to the higher-level listener stimulus directly, or indirectly via the trained intraverbal relations between lower-level names and the higher-level responses. If the intraverbals were used, this could have happened in two ways.

For example, on hearing /where's the noom?/ the child may have produced the intraverbal ("noom") "feb" "noom" "feb" and in responding as a listener to the last utterance in the intraverbal chain, select the correct alien in the same way as in the lower-level listener test.

Alternatively, on hearing /where's the noom?/ the child may have looked at the aliens on the table, one at a time, produced the intraverbal in relation to one of them (e.g., "feb-noom-feb-noom", when looking at the feb), possibly followed by producing the intraverbal relevant to the other alien while looking at it, and then select the alien for which (part of) the intraverbal matched the verbal prompt.

The vocalisations during the tests give some insight in this. Adam and Simon produced just the lower-level name (on one or two trials) in response to hearing the higher-level name, before or after selecting an alien. On another trial, Simon selects the correct alien, and then says, "Yes, that one because hib-noom". Cameron says on 10 trials "tor-zaag", and on two other trials (incorrectly) "the noom-tor", while on two more trials he only echoes the higher-level name. On all these trials he produces these responses before selecting an alien. Jon and Sasha produce the full intraverbals on one or more trials, before selecting an alien.

In short, intraverbals are produced in all but two trials with relevant vocalisations, whereas the remaining two trials may point to responding directly to the higher-level names (or to echoic responses to these). However, in the latter case, the boy had used intraverbals in earlier trials. Therefore, his echoic responses to the higher-level names may have been followed by producing the intraverbal covertly, also because on some of the earlier trials he whispered the intraverbals.

Another crucial question is whether the children can *tact* at the higher level. The fact that 4 out of 6 children failed the transfer of function test in part or in full is not unexpected, from the perspective of the naming account. Had the children passed a tact test at the higher level, then the naming account would predict an increased likelihood of the children showing transfer at both levels. However, the different name levels in the present studies do add complexity, and also inevitably lead to longer studies, and that may make it harder to control the factors that play a role in the verbal behaviour at issue. That is, the more complex the test task, the more component skills need to be trained, and the longer the learning process of these component skills, the harder it may be to predict test outcomes. Mike's full reversal of responses in the test of Study 2d is an example of that.

After novel response training and testing for transfer of function, Lowe et al. (2005) tested the children for listener behaviour to the experimenter's production of these novel responses. This was also done in the present studies; listener behaviour to the gestures was tested in Study 1f and Study 2f, while listener behaviour to the animal cries was tested in Study 1g and Study 2g. In all four studies, the children were tested with the aliens that had been involved in training of the production of the gestures and cries, and both participants passed these tests, as did the participants in Lowe et al.'s (2005) study. In Studies 2f and 2g, the test also involved the aliens that had *not* been involved in training, making it into an additional transfer of function test, but this time for listener behaviour. Mike, the only participant, showed behaviour that was similar to his behaviour in the test for transfer of the production of the novel responses (see the comparison earlier in this chapter). What was said earlier in the discussion of the transfer of function data regarding success and failure and the relation to the naming account, applies here as well.

The second way of investigating the functional properties of common name or category relations, listed by Lowe et al. (2005) along with transfer of function testing, was testing of category sorting among sets of arbitrary stimuli with common names. In the present studies, this was done in Study 1h and Study 2h.

Studies 1h and 2h were category match-to-sample tests, in which all eight aliens were lined up on the table. On each trial, the experimenter held up one of the aliens as the sample stimulus, and asked for "the others". Both Jon (in Study 1h) and Mike (in Study 2h) passed the category match-to-sample test (Test 1) at the lower level, by correctly selecting the other lower level category member, on a certain number of trials. For example, when Hib 1 was held up, the children selected Hib 2. This correct sorting, without being trained to do so, and without ever having seen the category members together, shows that teaching the children to tact the alien animals at the lower name level had established these arbitrary stimuli as a lower level class or category. Furthermore, Jon (in Study 1h) also passed the test at the higher level; that is, when presented with a zaag (e.g., Hib 1) as a sample stimulus, he correctly selected all the other zaags (Hib 2, Tor 1 and Tor 2), and the same for the nooms, over many trials. This indicates that the combination of tact training at the lower name level, and intraverbal training to link lower-level names and potential higher-level names, had also established the arbitrary stimuli as a higher-level class or category for him. Again, this categorisation came about without direct training, and without his ever having seen the category members together. Mike (in Study 2h) failed the test at the higher level.

The successful lower-level categorising after tact training is in line with Lowe et al. (2002), who also found in several experiments that teaching children to tact arbitrary stimuli established those stimuli as a class or category. Horne and Lowe

(1996) actually proposed that naming and not tacting would establish the classes, but as pointed out by Lowe et al. (2002), children of 2 to 4 years old usually learn listener behaviour in the course of tact training, which was confirmed by the Lowe et al. (2002, and 2005) studies, and in the present studies. As noted before, the Lowe et al. (2002, and 2005) studies did not involve names at different levels.

To be able to judge the data of Jon's successful and Mike's unsuccessful categorising at the higher name level, it would be useful to know whether the two boys could tact the stimuli at the higher level. This issue was also raised with regard to the transfer of function data. Both boys showed appropriate speaker and listener behaviour at the lower name level, and appropriate listener behaviour at the higher level prior to the category match-to-sample test, although Mike showed an increasing number of errors in his listener behaviour at the higher level in maintenance and review trials near the end of the experimental sessions. Whereas Jon's remark after a correct selection ("they're all zaags") suggests that he may also be able to tact the aliens at the higher name level, we do not know whether Mike's verbal repertoire would include tacting at the higher name level. If Mike was not able to tact the aliens at the higher level, then his failure to categorise at the higher level was to be expected from the perspective of the naming account, and from the Horne et al. (2004, and 2006) and Lowe et al. (2002, and 2005) studies; all these studies found that children who could tact arbitrary stimuli were able to categorise them, while children who showed appropriate listener behaviour to the stimuli but not tacting were not able to categorise. Some of the children in the latter studies who categorised correctly after tact training did need to tact the sample before making a selection, so they failed category match-to-sample Test 1 (which did not involve tacting the sample), but then went on to pass Test 2 (which did involve tacting the sample stimulus before selecting other category members).

Because Mike failed to categorise at the higher name level, he was also given category match-to-sample Test 2, and his performance in that test provided some information on whether tacting the stimuli at the higher level was part of his repertoire. Unfortunately, he only had six trials and then he refused to play any more. On all six trials, he correctly tacted the aliens at the lower level but he only selected the other lower level category member correctly on two of six trials. It should be noted here that between his correct tacting at the lower level and his category sorting there was interference of the second verbal stimulus prompting the child to tact the sample stimulus at the higher level. It is possible that he would have correctly selected the other lower level category member if given the chance to do so immediately after tacting the sample stimulus at the lower level. This is likely because he passed the earlier categorisation test at the lower level. When prompted to produce the correct tact at the higher level, he was only able to do so on two of the six trials he completed. On the other four trials he produced the intraverbal for the sample alien. And the verbal prompt "and is it a zaag or a noom?" seemed to puzzle him several times, even after producing the correct intraverbal. It was unfortunate that Mike refused to play any further, because further trials could have given us valuable information. It would have been interesting as well to see how Mike would have performed in another attempt at Test 2, if after his failure to categorise he could have had tact training at the higher level. The findings from the Horne et al. (2004, and 2006) and Lowe et al. (2002, and 2005) studies suggest that he might have passed Test 2 then.

As for naming (tacting and listener behaviour) during category match-to-sample Test 1, the naming account suggests that a child who successfully categorises the arbitrary stimuli, would tact the sample, either overtly or covertly, and hear the self-produced verbal stimulus, which would evoke the listener response of selecting the other category member(s). In this case, that would be tacting at both levels. Although Jon was not asked to tact the aliens at either level, and he did not do so *overtly* apart from the one trial in which he tacted the zaags after making a selection, there is a possibility that he did tact the aliens *covertly* before making his correct selections. Jon's comment "they're all zaags" suggests that he did indeed base his selection on the higher-level names.

Jon's categorisation was immediate, that is, his response latencies were not long, especially after the first few test trials. Throughout the lengthy test, he was clearly thinking carefully, while looking at the aliens, and giving full attention to the task. That is, he looked at each of the aliens, in silence and with full focus before, during, and after his selection. His gaze would only be broken when the selection was completed. All in all, his behaviour seemed like that of a chess player! Several times during the test, Experimenter 1 asked Jon whether he wanted to take a break from the test and go back to the other children, but he preferred to continue, and so all 48 trials were conducted in one session. Jon made only a few errors.

Perhaps the most important topic to be discussed in relation to the data is that of emergent behaviour. That is, behaviour that does not require direct reinforcement or training to be brought about. Emergent behaviour is acquired while a child learns other behaviour. In previous studies (e.g., Lowe et al., 2002, and 2005), the emergence of many new, untrained relations was reported. Likewise, in the present series of studies (1a-h, and 2a-h) several new relations between the arbitrary stimuli emerged, again without training or reinforcement. These include the following.

From tact training at the lower name level, listener behaviour emerged for all 26 children.

In addition, from intraverbal training, linking the lower-level names with the higher-level names, listener behaviour at the higher name level emerged for 13 out of 16 children (in Leg 1, and 6 out of 10 children in Leg 2).

When, after tact training and intraverbal training, the children were given function training (i.e., gesture training) to two aliens, one from one higher level category (the zaags) and one from the other (the nooms), 2 out of 6 children (Jon, and Simon) produced the correct gestures to the remaining six aliens that had not been involved in training. They thereby showed full transfer of function both at the lower, as well as the higher name level. Of the other 4 children, Alun and Huw showed partial transfer of function at the lower and higher level; they produced the correct gestures for one pair of aliens.

When 4 children (Jon, Huw, Mike, and Simon) were then given a new type of function training (training of animal cries) to two other aliens, again one zaag and one noom, Jon produced the correct gestures to two alien pairs that had not been involved in training, each consisting of a member of the same lower-level category and a member of the same higher-level category as the stimuli used in training. In a retest, Huw produced the correct gestures to one alien pair (one lower-level category member and one higher-level category member).

In a next test given to 2 children (Jon, and Mike), a new verbal prompt was shown to evoke both previously trained responses for Jon (though never on the same trial), while it evoked one of the two trained responses (consistently the animal cries) for Mike.

Then, for both Jon and Mike listener behaviour for the gestures and the animal cries was shown to be in place for the aliens involved in training of the production of these responses. Mike was also tested for transfer of this listener behaviour to the aliens that had not been used in training, and he showed the correct listener behaviour for one pair of aliens for the gestures (but for none of the alien pairs for the animal cries).

Finally, in the category match-to-sample test, when given one zaag (e.g., Hib 1) or one noom (e.g., Feb 2) as a sample stimulus, both of the tested children (Jon, and Mike) showed categorisation at the lower name level, as a result of the tact training in Study 1a; they correctly selected the other member in the same lower level category as the sample stimulus (e.g., Hib 2, or Feb 1) on a certain number of trials, without ever having seen the category members grouped together. Furthermore, the tact training combined with the intraverbal training in Study 1b, had also established the arbitrary stimuli as a class or category at the higher name level for Jon in Study 1h; he passed Test 1 at the higher level as well, also without ever having seen the category members grouped together. So he not only selected the correct lower-level category members (e.g., Hib 2, or Feb 1) but also the other higher-level category members (the two other zaags, for example Tor 1 and Tor 2, or the two other nooms, for example Lup 1 and Lup 2). Again, all these behaviours, or new relations between the arbitrary stimuli, emerged without training or reinforcement. These are interesting data in light of Chomsky's (1959) criticism of Skinner's (1957) book *Verbal Behavior*. In his review, Chomsky unjustly ascribes to Skinner the doctrine that "slow and careful shaping of verbal behavior through differential reinforcement is an absolute necessity" (1959, p. 42). As pointed out, the data summarized above list several examples of verbal behaviour that are acquired without reinforcement, and that can be accounted for within a behaviour analytic approach to verbal behaviour that builds on Skinner's book *Verbal Behavior*. Lowe et al. speak of "the extraordinary generative power of naming" (2005, p. 62). Lack of generativity was one of Chomsky's main problems with Skinner's behavioural account of verbal behaviour.

Some aspects of the data are puzzling. For example, why do some children show such persistent errors in tact training at the lower name level that they don't complete this training? For some children this could have to do with the learning criterion in tact training being very stringent. The criterion is 10/12 correct, and the probability of producing 10 or more per block correct by chance is minimal (p=.019). Furthermore, this criterion was extended to apply to two consecutive blocks, and even then, there was the additional restriction that the two errors that were allowed per block could not involve the same alien. All in all, these were very strict criteria, chosen to make absolutely sure that the child could tact, before administering the listener behaviour test, and before moving on to the follow-up studies that were all based on the child being able to tact the aliens at the lower name level. For children who consistently produced the same incorrect name for one or more particular alien(s), in line with the naming account one could imagine that when a child coincidently says the same incorrect name several times while looking at a particular alien animal (or familiar shape), the pairing of seeing that shape and hearing her own production of the name may then make it more likely that the child will say the same incorrect name again when seeing that particular alien. This can happen because "the sight of (that alien) becomes a frequent antecedent for the utterance (of the incorrect name); this establishes the object (the alien) as a discriminative stimulus for the child's (...) future utterance of (the incorrect name)" (Horne & Lowe, 1996, p. 199). If the child is then tested and shows corresponding listener behaviour, we know that an incorrect name relation has developed.

Something similar can happen in intraverbal training; an incorrect name link may coincidently be uttered, and when this is later repeated, it may strengthen the incorrect intraverbal relation. This could also be a reason why intraverbal training took so very long for some children, because making the same mistake over and over again had established a new intraverbal relation that was very hard to eliminate.

It is also puzzling that 4 children in Study 2b passed the listener behaviour test at the higher level for Leg 1, but then failed this test for Leg 2 of the hierarchy of names. For some children, there may have been an effect here of the long experimental procedure, also due to added pre-training. Or perhaps there was a switch from use of names to use of physical features, as the children's reference to physical features in the post-test interviews indicates. Researchers within psycholinguistics have proposed that children may have a *shape bias* in language learning (Jones, Smith & Landau, 1991); that is, that children primarily generalise names of objects on the basis of common shape, rather than, for example, colour or size. Smith (1995) found no shape bias in 18-month-old children, but a strong shape bias in 24 month olds, which would suggest that this shape bias would develop between the ages of 18 and 24 months. With regard to the data in the present studies, it remains puzzling why some children based their behaviour on names in the listener behaviour test for Leg 1, but then switched to using physical features in the test for Leg 2. Such a switch could perhaps have resulted (in part?) from the nursery animal theme, stimulating the children to focus on physical features of (and similarities and differences between) animals, along with a general emphasis of caregivers on physical features of objects in the environment of the developing child. It should be noted in relation to this, that the children's repertoire of verbal and nonverbal behaviour develops rapidly at this age, and mainly outside the experimental setting. Developments outside the experimental setting may very well have an effect on behaviour during the experimental sessions. And perhaps more than in any other type of behaviour, multiple causation plays a role in bringing about verbal behaviour (see Skinner, 1957), which is a complication for any researcher trying to account for verbal behaviour and its development.

A general focus on physical features in the interaction of caregivers with the child who is in the process of learning verbal behaviour, in combination with a shape bias, may also explain the next puzzling phenomenon; that of complete reversal of responses found in tests for listener behaviour at the higher name level, as well as transfer of function tests. Again, the children themselves explain their behaviour by reference to physical features of the aliens.

One child in Study 1a, Alun, explained during training how he used the shape of one of the aliens to remember its name. For him, Feb 2 was the long shape with spikes on the side. When this alien was targeted, Alun would sometimes say, "fff … feb". At some point he explained, "looks like a fork" and he added that he then thought, "f, for fork … f, for feb".

In addition to the above, some general issues should be raised. First of all, failure of the tests in Studies 2b, c, and h, cannot be attributed to the children's lack of understanding of the instructions, because in these studies the children had received pre-training and had been tested with familiar shapes, before moving on to the part of the studies in which the aliens were employed. In view of the fact that from Study 1c and Study 2c onwards performance in the first series of studies was roughly the same as, or better than, in the second series of studies, it is likely that failure in Studies 1 was not due to difficulty with the instructions either. Moreover, the behaviour of the children during pre-training in Studies 2 suggests that children of this age generally do not have difficulty with instructions of this kind.

As seen in Chapter 3 and Chapter 4, success and failure was not due to age either. In Study 1a, for example, three of the younger children have the smallest number of trials to criterion when learning to tact the aliens, and the picture is not very different in Study 2a. And Jon was the youngest child to participate in Study 1c (4 years old at testing), but the only one to show full transfer of function. The oldest child in Study 1c, Sasha, who was 5 years and 3 months at testing, did not show transfer to any of the pairs. Simon, the only one to show full transfer of function in Study 2c, was the oldest child in that study; he was 4 years and 6 months at testing. Alun, 4 years and 1 month at testing, showed transfer of function to one alien pair in Study 1c. Huw, who had similar results in Study 2c, was 3 years and 6 months old at testing.

What *did* influence failure and success, for example, was the child's on task behaviour. The experimental work extended over a long period of time[4], even more so in the second series of studies because of the added pretraining, and for most children that had a negative effect on how well they attended to the tasks, especially during training sessions, and sometimes during testing. Also, test sessions were

[4] Participation varied from one month to one and a half years.

necessarily long, due to review trials at the start, and large numbers of test trials. For several children (but by no means all) it was found that the longer the training procedures and the test sessions took, the more they displayed behaviours that were not beneficial to good task performance. The observed interfering behaviours were, for example, moving on the chair a lot, looking everywhere but to the aliens, talking (in length) about things irrelevant to the task, joking or being silly in other ways (like falling off the chair after every other trial), or just gazing. All these behaviours, together with asking regularly "are we done now?" or "are we reading a story now?", were signs of the children being less attentive to the experimental tasks. Children's on task behaviour, and consequently their performance, was negatively affected as well by heat, tiredness, toilet training, birthday parties and bouncy castles in the nursery, the Christmas season, and group activities the children were engaged in just before the session (e.g., dancing, running outside, or a long walk). These influences are unavoidable in research with children of this age, and the setting in which the research was conducted.

Finally, almost all children joked at times by producing the incorrect responses ... with a smile! Sometimes they even did this on crucial moments, when they were very close to reaching criterion, or during practice trials before a test.

To complete this section on the experimental data and the literature, let us now briefly consider how the three behaviour analytic approaches to categorisation described in Chapter 1 (stimulus equivalence theory, relational frame theory, and the naming account) would relate to the data presented in this thesis.

Stimulus equivalence theory (see Sidman, 2000) would predict that through lower-level tact training, and the intraverbal training that followed, two equivalence classes would be established.

For example, a child learns to tact two aliens as febs, and two other aliens as lups at the lower name level, and then learns the intraverbals Feb-Noom-Feb-Noom, and Lup-Noom-Lup-Noom (in Leg 1 and Leg 2 of the study, respectively). Note that, in effect, the intraverbals are combinations of speaker and listener elements. To illustrate this for one intraverbal, producing the intraverbal "Feb-Noom-Feb-Noom" can be written as: say "feb", hear /feb/, say "noom", hear /noom/, say "feb", hear /feb/, say "noom", hear /noom/. After tact training and this intraverbal training (of the Feb-Noom, and Lup-Noom links), stimulus equivalence theory would predict the establishment of a class consisting of the following stimuli and responses:

alien Feb 1	alien Feb 2	"Feb" (say name)		/Feb/ (hear name)	
alien Lup 1	alien Lup 2	"Lup"	/Lup/	"noom"	/noom/

Through tact training and intraverbal training, a similar, separate class would be established for the hibs and the tors but then including "zaag", and /zaag/. With all these events, as just listed, being part of the same class, any novel behaviour trained to one of these, would be expected to transfer to all the other events, without training. And apart from perfect transfer of function, also perfect category sorting would be expected. However, only 2 out of 6 children showed perfect transfer of function, and 1 out of 2 showed perfect category sorting.

For relational frame theory, the most relevant aspect of the present studies would be the pre-training with familiar stimuli that was added in the second series of studies. This provided the children with multiple exemplar training in preparation of both training and testing with the aliens. RFT would predict that children who had multiple exemplar training with familiar stimuli, would perform better on the test tasks than children who had not had such training, especially when it comes to the transfer of function and the category sorting tests. However, this is not what was found.

The added pretraining speeded up the learning process for lower-level tacting, although on average it had the opposite effect on the learning of the intraverbals with alien names. In addition, after pretraining, more children passed the listener behaviour test at the higher name level for Leg 1, but not for Leg 2. However, more importantly, pretraining did not result in better performance on the transfer of function and the categorisation tests. The transfer of function data in Studies 1 and 2 are very similar, and in the category match-to-sample test the child in Study 1h, who had no pretraining, showed near perfect performance at both the lower- and the higher name level, whereas the child in Study 2h, who had had pretraining, only passed the categorisation test at the lower name level.

The naming account, based on Skinnerian principles, predicts that children will pass the listener behaviour test at the lower level after tact training, as explained earlier in this section (and see Lowe, et al., 2002, 2005). Furthermore, it also predicts positive outcomes for the listener test at the higher level for Leg 1 of the hierarchy. The intraverbal training establishes a bidirectional link between the lower-level tacts and the potential higher-level names, such that hearing the higher-level label during the listener behaviour test should occasion saying and hearing of the corresponding lower-level tact. Through mediation of this lower-level tact, correct listener responses to the higher order label are brought about.

However, training of the intraverbals of Leg 2 may result in interference between the trained intraverbal chains. The verbal stimulus /Where's the Noom?/ may now evoke saying "Feb", or "Lup", or both, possibly in combination with the higher-level label. Therefore, the probability that the intraverbal will generate a lower-level tact (and corresponding listener stimulus) relevant to the two aliens presented is reduced. This may affect performance on the listener test at the higher name level for Leg 2.

Continued testing for lower-level tacting when the children were learning the intraverbals, may also have affected the performance on this listener test. These test trials for tacting served to maintain the discriminations between the Febs and the Lups. Naming the aliens at the higher-level would break down that discrimination, and lead to failure on the test. Therefore, there was constant extinction of saying the higher-level names in the presence of the aliens. As a result, the probability of emitting the higher-level names, or the intraverbals, in the presence of the aliens (that is, in the lower-level tact context) was much lower than in the intraverbal context. This should make the development of "Noom" and "Zaag" as name relations unlikely. And as a consequence, transfer of novel behaviour as well as category sorting should be largely confined to the lower name level.

That some children show partial transfer of novel behaviour at the lower as well as the higher-level (e.g., for one of the lower-level category members rather than both, and to one or two higher-level category members) may be explained in terms of the children formulating their own higher order names that are not in agreement with the experimentally defined higher-level categories. Although these were arbitrary stimuli, these names could be based on physical features. For example, the child could call some aliens "spiky", and others "straight" (as seen with Mike).

These self-formulated names would have been a competing source of control, and could explain why children failed transfer of function tests but were completely consistent in their responses to the particular aliens presented (see Horne et al., 2004).

In view of the above, the data seem most consistent with the naming account and Skinnerian principles, as opposed to stimulus equivalence theory, or relational frame theory.

Methodological and conceptual matters

In this section, several matters will be discussed that concern the methodology of the studies described in the previous two chapters.

The first conceptual matter is related to the training procedure in Studies 1a and 2a. This is referred to as a tact training procedure. Once each of the four animals in one set are introduced, the tact training procedure entails the child on each trial being asked for a targeted animal, "What's this?" When the child responds, feedback is given. So in addition to the object being present, the experimenter emits a mand for verbal action: that is, the verbal action of the child saying the name of the targeted alien animal. Now, Greer and Ross (2004, p. 144-145) maintain that the child's tact in a situation like this is a case of an *impure* tact, because of the verbal antecedent "What's this?" They write that one can only speak of a tact when the child says the name in the presence of the object, and there are no verbal antecedents like "What is this?" However, Skinner's (1957, p. 84) example of tacting is exactly like this, that is, with the mand for verbal action: in the presence of a red object, and possibly the mand "What colour is that?" the speaker produces the verbal response "Red". Elsewhere, he gives an example of an impure tact; someone has cooked a meal and when done, she says, “Dinner is ready”. Skinner states that this could be a pure tact, but when it is called out through the house and as a result the family comes running to the dining table, it is a combination of a tact and a mand. That is, there is a mixture of controlling relations characteristic of both tact and mand. In this case, Skinner (1957, p. 151) speaks of an impure tact. In line with this, Sundberg, San Juan, Dawdy, and Argüelles (1990), to name just one example, also speak of tact training when referring to a procedure that includes the verbal prompt “What is this?” in the presence of the object to be tacted.

A methodological issue is, that one could wonder why the higer-level names were not trained directly, in the presence of the aliens. In these studies, an attempt was made to determine whether it would be enough for children to hear lower-level and higher-level names together, to establish higher-level naming of objects or animals. It was anticipated that training of two names for the same animal might be extremely time-consuming. The results of Studies 1a and 2a show that teaching one name for each alien can already be very time-consuming. Teaching the children the intraverbal links was expected to be a quick, alternative way that was also more interesting theoretically, because in this way one could test for possible generative outcomes of intraverbal training (generative outcomes were indeed found for 6 children in Study 1b and Study 2b together).

Another conceptual and methodological issue concerns, for example, the names that are used. A question that could be raised in relation to these names is: how do you know whether a child interprets the names as common nouns, and not as proper names?[5] The animals are introduced one by one by saying, "This is a hib (feb/tor/lup)", suggesting a common name. It may be commented then that some proper names start with a schwa[6], for example Amelia, but then one could also ask how a child knows she is dealing with common names when she is told that something is "a table", "a dog" or "a cat". So this seems to be relevant not just to the names in the present studies, but also to names in real life.

In the present studies, there are several clues that the children do not seem to have a problem here. One clue is that after the introduction of each alien animal, the verbal prompt on each tact training trial is, "What's this?" and the children sometimes reply, "*a* hib (feb/tor/lup)", but sometimes they simply say, "hib (feb/tor/lup)". Furthermore, when a child has learned to tact the aliens in Set 1, a second set of four aliens is presented, also consisting of a hib, a feb, a tor, and a lup, which should make it more likely that the children interpret the names as common names rather than proper names, if they had not interpreted them as such already. When the second set was introduced to Ginny, she commented: "a different lup" and "a different hib". Some of the other children made similar comments (e.g., Kyle). Finally, all children passed the listener behaviour test at the lower name level, suggesting that none of the children had a problem with the verbal prompt in the test, "which one is *the* hib (feb/tor/lup)?"

Relevant to this issue is a study by Hall and Bélanger (2005) on use of "range-of-reference information" in word learning by 3-year-olds. Hall and Bélanger presented the children with a new word, DAXY, in a particular play situation with three identical toy rabbits. The word was presented with, or without linguistic cues.

---

[5] An example of a common noun is "table", and an example of a proper name is "John".
[6] The schwa is the sound "a" as is it produced in "a bed".

A cue could suggest that the new word was a proper name ("Look! His name is DAXY"), or an adjective ("Look! He is very DAXY"). It was found that when they presented the new word without any specific linguistic cues regarding whether it was a proper name or an adjective (by saying, "Look! DAXY"), children would react to the word as to a proper name, if the experimenter had applied it to only one object. That is, in a listener behaviour test ("Now look! Which one is DAXY?"), they would select the rabbit that had been called DAXY by the experimenter, but not an identical rabbit in a different position. However, if the experimenter had applied the new word to two of the three identical objects, the child's listener behaviour would extend to include the third identical object as "daxy", thereby reacting to the word as to a common noun or an adjective. Hall and Bélanger conclude that the children used the range-of-reference information to guide their behaviour.

When applying this to Study 1a and Study 2a in this thesis, these findings suggest that presenting one alien as "a hib", and then introducing a second, different looking, alien with the same name would result in the children responding to "hib" as to a common noun, if introducing the aliens by saying, "This one is *a* hib" would not be enough to establish this.

Another question that could come up in the course of reading this thesis is: why speak of lower- and higher-level names, rather than basic and super-ordinate names? The reason for this is that the lower and higher name levels focussed on in the present studies don't map onto the distinction in the cognitive developmental literature between basic and super-ordinate levels. To illustrate, examples of basic level names would be "dog" or "cat", and a super-ordinate level name would be "animal". In the present studies, the lower-level names "hib", "feb", "tor", and "lup" may seem comparable with the basic level, but basic level entities have physical commonalities, whereas the hibs, for example, do not. In addition, the hibs, febs, tors, and lups are all "animals" at the superordinate level. Therefore, the names "zaag" and "noom" should be seen as names at an intermediate level, perhaps comparable with dogs and cats being called mammals.

One final topic in relation to names is that the animal cries "Boo!" and "Raagh!" have the same vowels as the higher level names noom and zaag. The studies in which each of these animal cries were trained to one zaag and one noom, had 4 participants (1 in Study 1d, and 3 in Study 2d). For Jon (in Study 1d) and for Huw (in Study 2d), the "Boo!" cry was linked to the noom, and the "Raagh!" cry was linked to the zaag. This could have helped these children to perform better in the transfer of function test than the 2 children (Simon and Mike) for whom the "Boo!" cry was linked to the zaag, and the "Raagh!" cry to the noom.

However, none of the children did very well in the initial test; Jon reached criterion for one of the pairs, whereas Huw, Mike and Simon reached criterion for none of them. In a retest, Jon reached criterion for two pairs, Huw for one, Simon still for none (mostly because he reverted to the Raagh! response on many trials). Although Jon and Huw did best in these studies, Jon's performance was not as good as it was in Study 1c, and Huw did not perform well either.

If their transfer of function responses were due to the correspondence of vowels in the higher-level names and the animal cries, then why did they not show transfer for all three pairs, as Jon did in Study 1c? Mike's complete reversal could have been due to him incorrectly producing the Boo! cry for nooms, and the Raagh! cry for zaags. But Huw, for whom there *was* vowel correspondence, also showed reversal of responses for four out of six aliens (and he showed a tendency to reverse for another alien, for which he had 3/10 correct). Furthermore, in Study 2c when tested for transfer of the trained gestures, where there was of course no issue of sound correspondence, Mike also showed reversal for one alien, and a tendency to reverse for another three aliens. And in Study 2b, reversal of responses was found for some children as well, in the listener behaviour test at the higher name level. Finally, in his post-test interview Mike referred to physical features to explain his responses. To conclude, it is not at all clear whether there was any influence on the children's performance of the vowel correspondence between names and animal cries.

With regard to the gestures that were trained to aliens in Studies 1c and 2c, it could be commented that it is very implausible that the aliens could make these gestures (comparable to greetings), because the gestures are produced with hands, and the aliens don't have hands! However, to the children this did not seem to be a problem. Only one child, Simon, noticed this. When he was taught that Feb 2 makes the "hands in front" gesture, he asked: "How does that one do that?"

Simon was also the only child who asked questions about the word game (see Appendix C). He asked the experimenter what "zaag" meant. When the experimenter said "Oh, it's just a name", he asked why different looking aliens had the same name. Later in maintenance training for tacting at the lower level, Simon correctly tacted Lup 1, and added: "It's a zaag, zaag is the second name for lup."

Another matter to be discussed here is that the procedures that were employed in the present studies were very different from those used in applied settings by behaviour analysts. In applied settings, the procedure for teaching tacts, for example, is tailored to the individual child, and changed in reaction to the child's success or lack of it. It is unthinkable that a child would have dozens or even hundreds of trials (especially relevant to Study 2b) without much improvement, and yet no changes to the procedure, such as reducing the number of stimuli presented, which was done in Study 1a with Davey, Wendy and Ellie, only after an extended time of no improvement. However, in experimental settings uniformity of procedure is a must, to ensure that the data are comparable across children. The role of reinforcers is discussed below.

Future directions

In this section possible future directions for research are explored. Replication of the studies described in the previous chapters would be desirable, especially the transfer of function studies and the category match-to-sample tests, where the number of participants was very small. Of course, a complicating factor in this, is that to be able to participate in these later studies, children will need to undergo tact training and intraverbal training first and testing for listener behaviour at the lower and higher level, and in addition, they will need to pass these tests. Therefore, it is difficult to increase the number of participants in the later studies. However, if replication were to take place, several recommendations can be given for improvement or extension of the studies.

First, the number of trials needed in pre-training for Study 2a could be reduced for many children by taking out the fruits and vegetables, and replacing them by shapes that the children can tact more readily. Many children at this age (3 - 4.5 years old) have difficulty distinguishing fruits from vegetables, and therefore need a considerable number of trials to reach criterion for both sets of familiar shapes in pre-training. But it is not easy to find shapes to replace them with, because of the requirement that the same higher-level name should apply to these shapes, and that the lower-level names should apply to a class of objects or entities that look very different from each other. Perhaps the two vegetables could be replaced by two sweets, such as a lolly, and star shaped liquorice.

In Studies 1b and 2b, to ensure comprehension of the instructions in training and testing, it would be enough to give pre-training for one leg of the hierarchy, and then to test the children for listener behaviour at the higher name level to the familiar shapes involved. In combination with taking out tongue breaker (that is, hard to pronounce) intraverbals such as "fruit-food-fruit-food" (and perhaps "reptile-animal-reptile-animal"), this could reduce the number of trials in pre-training considerably. Further complications in the present studies were that some children incorrectly produced the not unlogical intraverbal "fish-food-fish-food", and that some children confused pre-training with alien training in some way, because one of the higher-level names for the familiar shapes is "animal", while the aliens are animals too. For example, in the alien name word game, Tomos said, "feb-animal-feb-animal", while Louise said, "hib-animal-hib-animal". When Louise was corrected by the experimenter telling her it was "hib-zaag", she said: A hib is also an animal, isn't it?"

Before function training in Studies 1c and 2c, an additional tact test at the higher name level could be given to children who passed the listener behaviour test at the higher level in Studies 1b and 2b. That could possibly yield two groups of participants, one group that showed only appropriate listener behaviour at the higher name level, and perhaps a second group of children in whom not just listener behaviour but also tacting at the higher name level is in place. If this would indeed result in two groups of children, then it would be very interesting to compare how both groups perform in the transfer of function test. The naming account would predict that children who show correct listener behaviour but not tacting, would fail subsequent transfer of function and category match-to-sample tests, while children who can tact the aliens at the higher level would pass these tests. If the "listeners" would indeed fail the transfer of function test, then they could be given tact training at the higher name level, followed by a repeat transfer of function test.

In the transfer of function tests, it would be useful to take apart the two levels, by first presenting the two aliens in the same lower-level category as the aliens that had been part of function training, and then the higher-level category members. This would also eliminate the possibility of reaching criterion for a higher-level category member through selection by exclusion. That is, when a child shows transfer of function at the lower name level, it is possible to then infer the correct response for the higher-level category member that is presented with this lower-level category member. Selection by exclusion was suspected in a few children, and supported by vocalisations during testing by Jon. If the two levels were separated in the test, the children might be more likely to show full transfer of function at the lower level, without the complicating factor of a lower-level category member being presented with a higher-level category member that might be targeted in the first trial, which in turn could influence results for lower-level transfer of function. In case of failure of the test at the higher level, or at both levels, function training could be given in which the child would be required to tact the aliens at both levels before producing the relevant gesture (or the animal cry). In testing, the child could then be prompted to tact at both levels as well, before producing a gesture or a cry. In the present studies, this incorporating of tacting was done with Sasha, who then performed much better on the transfer of function test.

As an alternative to giving the children an additional tact test at the higher name level before starting function training in Study 1c or 2c, this tact test at the higher level could be given after a child has passed category match-to-sample Test 1.

If a child fails category match-to-sample Test 1, the two levels could be taken apart again, as suggested for the transfer of function test. This could be done by giving the child a Test 2a, which could consist of a lower level category match-to-sample test with tacting at the lower level before responding to a lower level prompt, "Can you give me the other one like this one?" If the child would pass this test, a Test 2b could be given, which would be the full category match-to-sample Test 2 as described in Study 2h, requiring tacting at both levels followed by a prompt for selection of category members at both levels, "Can you give me the others like this one?"

As for the selection prompt used in the category match-to-sample tests as described in the previous two chapters, it might be better to replace "Can you give me the others like this one?" with "Can you give me the others?" to avoid stimulating the children to select others that *look like* the sample alien.

Apart from replication of the studies with the additions described above, other experimental options could be explored.

As reported, differences in performance were found between boys and girls. It was also found that many of the boys in these studies were more engaged in the sessions with the aliens than most of the girls were, especially when the experimental procedures took longer to complete. And although all children were initially keen to play with the new aliens, some boys were clearly very excited at the sight of the aliens, and couldn't wait to find out more about them. For example, Sebastian's first reaction to the aliens was, "Are they dinosaurs? What are they? Are they dinosaurs?" The differences in performance could very well be related to their choice of toys outside the experimental sessions. From casual observation of the children it was found that despite a policy of equal opportunities in Tir na n'Og Nursery, which meant that all children had free access to all the toys, there remained a clear gender bias in the toy selection when given a free choice. Most of the girls played with pink and purple fairies and Barbie dolls, and made ballet moves or watched a ballet video, while many of the boys played with dinosaurs, building blocks, police cars and fire engines. In light of these differences, and the results of the present studies, it might be interesting to study gender differences in trials to criterion if the aliens were to be made fluffy and pink! This might influence the experimental findings.

Further experimental options are to make colourful picture story books featuring the aliens that could be read to the children, to develop jigsaws and board games also featuring the aliens, or to have the aliens appear as cartoon figures in video clips and/or computer games, or in a trial and error computer learning program.

These methods and media could all be used when teaching children to tact the aliens, and could make the learning situation resemble the present day real life situation of children learning verbal behaviour more. There seems to be a general assumption among parents of young children that their children learn very quickly, and perhaps most impressively when learning things from television or the computer. This could be tested by use of the aliens, and then comparing the results to the results of the present studies.

The various methods and media could be used separately or combined, for different groups of children, to see what is the most effective way to teach children names. It may be that variation in presentation affects how fast children learn, and more than anything, variation in presentation could lead to more on task behaviour, and less distraction. One parent commented that her 4-year-old son could not be distracted when watching television, so perhaps use of that medium in learning could be useful. One could make sure that, for example in the first story, video clip, or computer game, each of the four aliens in a set would be introduced at the first appearance (the name would be provided), while after that the child would have to provide the name of each of them twice in the first play session. In the following sessions, the child would be prompted for each name three times in one story, clip or game. That way the sessions would resemble the trial blocks in the present studies, at least in the number of times each alien is targeted, which would make the data more comparable. A computer program could be developed to test the children for listener behaviour, in which the child can touch the alien (via a touch screen method) when a name is heard. Results from these studies could be used in teaching practices with normally developing children, and children who show a developmental delay.

Whichever of these experimental options is explored, there should be careful consideration of the reinforcers to be used. In the present studies, during a training session there was verbal praise for correct responses, and at the end of each session the child selected a story to be read by the experimenter. The verbal praise and the story served as rewards, but the question is: were they also reinforcers? That is, did they bring about an increase in the number of correct responses produced by a child?

Judging from the facial expressions of the children (eyes and face lighting up, smiling) these rewards were definitely pleasurable throughout the studies. And judging from the steady or sometimes large increases in the number of correct responses at the initial stages of the studies, they also served as reinforcers to begin with. However, several children reached a stage where their number of correct responses remained stable. For most of them, this only lasted a few trial blocks, but for others this was the case over many trial blocks (for Davey and Wendy, for example).

When there was no increase over many trial blocks, occasionally a music book was used during the session on which a child could press buttons, one for each, or for every three, correct responses. If the schedule was one button for every three correct responses, the experimenter held up one finger for each correct response, and when three fingers were held up, the music book was presented. This often worked as a strong reinforcer, but for some children it lost it's reinforcing effect quickly (Ellie, for example).

Near the end of all experimental work, magic cards were introduced. The children could put stickers on them, received at the end of a session. When the magic card was full, the child could take the card home. For some children this worked as a very strong reinforcer, for others the reinforcing effect wore off quickly, and for yet others it had hardly any reinforcing effect in the first place. Furthermore, for some children particular rewards did not work as a reinforcer with regard to the number of trials correct, but it did increase the amount of time the child showed on task behaviour.

Altogether this shows how rewards can differ in their reinforcing effect. What works as a reinforcer for one child, may not work as such for another child. And it also illustrates that for one child a particular reward may have a reinforcing effect at one time but not at another time. Therefore, in future studies it would be best to provide several different types of rewards, such as books, toys, stickers, music books, or games. A preference assessment could be done (regularly), and a token economy set up. Also, to make sure that the toys that are used are not just rewards, but also reinforcers, perhaps a pilot session could be done in which each child would be given a few learning tasks, and different rewards used for comparable tasks, to measure the effect of reinforcers.

In an attempt to avoid a reinforcing effect wearing off, it would desirable to vary the collection of rewards presented to the child over time, and to add new ones regularly. However, financial restrictions (especially within PhD research) will limit the possibilities for creating this ideal situation. And a further complication is the longitudinal nature of these studies; sometimes studies like these take several months to complete. Because the task remains exactly the same throughout these months, it is almost impossible to avoid the effects of specific reinforcers wearing off. If anything is clear here, it will be that the matter of reinforcers is a complex one.

References

Andresen, J. T. (1990). Skinner and Chomsky thirty years later. *Historiographia Linguistica, XVII (1/2)*, 145–165.

Andresen, J. T. (1992). The behaviorist turn in recent theories of language. *Behaviorism, 20*, 1–19.

Anisfeld, M. (1984). *Language development from birth to three*. London: Erlbaum.

Barnes, D., McCullagh, P.D., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *The Analysis of Verbal Behavior, 8*, 19–30.

Barnes-Holmes, D., Barnes-Holmes, Y., Smeets, P. M., Cullinan, V., & Leader, G. (2004). Relational frame theory and stimulus equivalence: Conceptual and procedural issues. *International Journal of Psychology and Psychological Therapy, 4*, 181–214.

Blackledge, J. T. (2003). An introduction to relational frame theory: Basics and applications. *The Behavior Analyst Today, 3*, 421–433.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Catania, A. C. (1988). The operant behaviorism of B. F. Skinner. In A. C. Catania & S. Harnad (Eds.), *The selection of behavior* (pp. 3–8). Cambridge, USA: Cambridge University Press.

Catania, A. C. (1998). *Learning (4th ed.)*. Upper Saddle River, NJ: Prentice Hall.

Chomsky, N. (1959). Review of B.F. Skinner's *Verbal Behavior. Language, 35,* 26–58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Czubaroff, J. (1988). Criticism and response in the Skinner controversies. *Journal of the Experimental Analysis of Behavior, 49*, 321–329.

Czubaroff, J. (1989). The deliberate character of strategic scientific debates. In H. W. Simons (Ed.), *Rhetoric in the human sciences* (pp. 28–47). London: Sage.

Dale, R. (2004). Cognitive and behavioral approaches to language acquisition: Conceptual and empirical intersections. *The Behavior Analyst Today, 5*, 336–358.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46*, 243–257.

Dickins, D. W., Bentall, R. P., & Smith, A. B. (1993). The role of individual stimulus names in the emergence of equivalence relations: The effects of paired-associates training between names. *The Psychological Record, 43*, 713–724.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice* (pp. 115–138). Hove, England: Erlbaum.

Dulany, Jr., D. E. (1959). Verbal Behavior. *Science, 129*, 143–144.

Eikeseth, S., & Smith, T. (1992). The development of functional and equivalence classes in high-functioning autistic children: The role of naming. *Journal of the Experimental Analysis of Behaviour, 58*, 123–133.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford: Oxford University Press.

Goldiamond, I. (1962). Perception. In A. J. Bachrach (Ed.), *Experimental foundations of clinical psychology* (pp. 280–340). New York: Basic Books.

Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation, 95*, 260–270.

Greer, R. D., & Keohane, D. D. (2005). The evolution of verbal behavior in children. *Behavioral Development Bulletin, 1,* 31–47.

Greer, R. D., & Ross, D. E. (2004). Verbal behavior analysis: A program of research in the induction and expansion of complex verbal behavior. *Journal of Early and Intensive Behavior Intervention, 1*, 141–156.

Hall, D. G., & Bélanger, J. (2005). Young children's use of range-of-reference information in word learning. *Developmental Science, 8*, 8–15.

Halliday, M. A. K. (1975). *Learning how to mean*. London: Edward Arnold.

Harlow, H. (1959). Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science (Vol. 2)*. New York: McGraw-Hill.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Reno, NV: Context Press.

Hayes, S. C. (1994). Relational frame theory: A functional approach to verbal events. In S. C. Hayes & L. J. Hayes (Eds.), *Behavior analysis of language and cognition* (pp. 9–29). Reno, NV: Context Press.

Hayes, S. C., & Barnes-Holmes, D. (2004). Relational operants: Processes and implications: A response to Palmer's review of Relational Frame Theory. *Journal of the Experimental Analysis of Behaviour, 82*, 213–224.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A post Skinnerian account of human language and cognition.* New York: Kluwer Academic/Plenum.

Hayes, S. C. & Berens, N. M. (2004). Why relational frame theory alters the relationship between basic and applied behavioral psychology. *International Journal of Psychology and Psychological Therapy, 4*, 341–353.

Hayes, S. C. & Hayes, L. J. (1989). The verbal action of the listener as a basis of rule-governance. In S. C. Hayes (Ed.), *Rule governed behavior: Cognition, contingencies, and instructional control* (pp. 153–190). New York: Plenum Press.

Hayes, S. C. & Hayes, L. J. (1992). Verbal relations, cognition, and the evolution of behavior analysis. *American Psychologist, 47,* 1383–1395.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 185–241.

Horne, P. J., & Lowe, C. F. (1997). Toward a theory of verbal behavior. *Journal of the Experimental Analysis of Behavior, 68*, 271–296.

Horne, P. J., & Lowe, C. F. (2000). Putting the naming account to the test: Preview of an experimental program. In J. C. Leslie & D. Blackman (Eds.), *Experimental and applied analysis of human behavior* (pp. 127–148). Reno, NV: Context Press.

Horne, P. J., Hughes, J.C., & Lowe, C. F. (2006). Naming and categorization in young children: IV. Listener behavior training and transfer of function. *Journal of the Experimental Analysis of Behavior, 85*, 247–273.

Horne, P. J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young children: II. Listener behavior training. *Journal of the Experimental Analysis of Behavior, 81*, 267–288.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*, 499–516.

Julià, P. (1983). *Explanatory models in linguistics: A behavioral perspective.* Princeton, NJ: Princeton University Press.

Knapp, T. J. (1992). *Verbal Behavior*: The other reviews. *The Analysis of Verbal Behavior, 10*, 87–95.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*, 407–414.

Leigland, S. (2001). Toward a science of verbal behavior: Progress and challenges. *The Behavior Analyst Today, 2,* 166–169. Available from: http://www.behavior-analyst-today.com/

Leigland, S. (2007). Fifty years later: Comments on the further development of a science of verbal behavior. *The Behavior Analyst Today, 8,* 336–346. Available from: http://www.behavior-analyst-today.com/

Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology, 56*, 201–239.

Lowe, C. F., & Beasty, A. (1987). Language and the emergence of equivalence relations: A developmental study. *Bulletin of the British Psychological Society, 40*, A42.

Lowe, C. F., & Horne, P. J. (1996). Reflections on naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 315–340.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorisation in young children: Vocal tact training. *Journal of the Experimental Analysis of Behaviour, 78,* 527–549.

Lowe, C. F. Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III. Vocal tact training and transfer of function. *Journal of the Experimental Analysis of Behavior, 83,* 47–65.

Luciano, C., Gómez Becerra, I., & Rodríguez Valverde, M. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior, 87,* 349–365.

MacCorquodale, K. (1969). B. F. Skinner's Verbal Behavior: A retrospective appreciation. *Journal of the Experimental Analysis of Behavior, 12*, 831–841.

MacCorquodale, K. (1970). On Chomsky's review of Skinner's Verbal Behavior. *Journal of the Experimental Analysis of Behavior, 13*, 83–99.

Mandell, C., & Sheen, V. (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioral Processes, 32*, 29–46.

Mead, G. H. (1934). *Mind, self and society.* Chicago: University of Chicago Press.

Moerk, E. L. (1983). *The mother of Eve—as a first language teacher.* Norwood, NJ: Ablex.

Moerk, E. L. (1990). Three-term contingency patterns in mother-child verbal interactions during first-language acquisition. *Journal of the Experimental Analysis of Behavior, 54*, 293–305.

Moerk, E. L. (1992). *A first language: Taught and learned.* Baltimore, MD: Paul H. Brookes.

Moerk, E. L. (2000). *The guided acquisition of first language skills.* Norwood, NJ: Ablex.

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development, 21*, 373–377.

Morris, C. (1958). Verbal Behavior, by B. F. Skinner. *Contemporary Psychology, 3*, 212–214.

Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization [Emory Symposia in Cognition]*, (pp. 11–24). New York: Cambridge University Press.

Nelson, K. E. & Bonvillian, J. D. (1973). Concepts and words in the 2-year-old: Acquiring of concept names under controlled conditions. *Cognition, 2*, 435–450.

Oah, S., & Dickinson, A. M. (1989). A review of empirical studies of verbal behavior. *The Analysis of Verbal Behavior, 7*, 53–68.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development, 59*, 441–449.

Osgood, C. E. (1958). Verbal Behavior, by B. F. Skinner. *Contemporary Psychology, 3*, 209–212.

Palmer, D. C. (1996). Achieving parity: The role of automatic reinforcement. *Journal of the Experimental Analysis of Behavior, 65*, 289–290.

Palmer, D. C. (2000a). Chomsky's nativism: A critical review. *The Analysis of Verbal Behavior, 17*, 39–50.

Palmer, D. C. (2000b). Chomsky's nativism reconsidered. *The Analysis of Verbal Behavior, 17*, 51–56.

Palmer, D. C. (2004). Data in search of a principle: A review of *Relational Frame Theory: A post-Skinnerian account of human language and cognition. Journal of the Experimental Analysis of Behavior, 81*, 189–204.

Poulson, C. L., & Kymissis, E. (1988). Generalized imitation in infants. *Journal of Experimental Child Psychology, 46*, 324–336.

Poulson, C. L., Kymissis, E., Reeve, K. F., Andreatos, M., & Reeve, L. (1991). Generalized vocal imitation in infants. *Journal of Experimental Child Psychology, 51*, 267–279.

Rakison, D. H. & Oakes, L. M. (2003). *Early category and concept development: Making sense of the blooming, buzzing confusion*. Oxford, England: Oxford University Press.

Richelle, M. (1976). Formal analysis and functional analysis of verbal behavior: Notes on the debate between Chomsky and Skinner. *Behaviorism, 4*, 209–221.

Richelle, M. N. (1993). *B. F. Skinner: A reappraisal.* Hove, East Sussex, UK: Lawrence Erlbaum Associates.

Schlinger, H. D. (1995). *A behavior analytic view of child development.* New York/London: Plenum Press.

Schoneberger, T. (2000). A departure from cognitivism: Implications of Chomsky's *second* revolution in linguistics. *The Analysis of Verbal Behavior, 17*, 57–73.

Schrier, A. M., & Thompson, C. R. (1980). Conditional discrimination learning: A critique and amplification. *Journal of Experimental Analysis of Behavior, 33*, 291-298.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion *(Zalophus californianus)* is capable of forming equivalence relations. *The Psychological Record, 43*, 823–839.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14,* 5–13.

Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 93–114). Hillsdale, NJ: Erlbaum.

Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.

Sidman, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127–146.

Sidman, M., Willson-Morris, M. and Kirk, B. (1986). Matching to sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6,* 1–19.

Skinner, B. F. (1957). *Verbal Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not inductor. In C. A. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development: The Minnesota Symposia on Child Psychology* (Vol. 28, pp. 1–32). Mahwah, NJ: Lawrence Erlbaum Associates.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language, 4*, 1–22.

Snow, C. E., & Fergusen, C. E. (1977). *Talking to children: Language input and acquisition.* Cambridge, USA: Cambridge University Press.

Stemmer, N. (1990). Skinner's Verbal Behavior, Chomsky's review, and mentalism. *Journal of the Experimental Analysis of Behavior, 54*, 307–315.

Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25*, 698–724.

Sundberg, M. L., San Juan, B., Dawdy, M., & Argüelles, M. (1990). The acquisition of tacts, mands, and intraverbals by individuals with traumatic brain injury. *The Analysis of Verbal Behavior, 8*, 83–99.

Terrace, H. S. (1985). In the beginning was the "name". *American Psychologist, 40*, 1011–1028.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical style. *First Language, 4*, 197–212.

Tomanari, G. Y., Sidman, M, Rubio, A. R., & Dube, W. V. (2006). Equivalence classes with requirements for short response latencies. *Journal of the Experimental Analysis of Behavior, 85*, 349–369.

APPENDIX A: CONSENT LETTER AND FORM IN ENGLISH AND WELSH

Dear Parent/Guardian,

At Tir na n-Og we are about to begin another study of language development in children which builds upon previous studies we have conducted here. The research mainly involves teaching the children the names of different objects, within a 'play' context. We are interested in finding ways to teach language effectively and hope to observe how this benefits the child's learning generally.

Every effort will be made to ensure that the process, in which our staff will name various objects for the child and assess his/her understanding of these names, will be very enjoyable for the children. The procedures to be employed have been used in a range of similar studies and the children have enjoyed their participation very much. Many of the parents were very positive indeed about the value of these daily one-to-one interactions.

This particular study is to be directed by Marleen Adema, a postgraduate member of our research team, working under my supervision. Marleen has experience in conducting infant studies in the Nursery. She will be happy to discuss the details of the study with you and will keep you fully informed as to how it proceeds.

It is difficult to determine how long it will take to observe the language effects we are investigating, but we anticipate that the study will last approximately 2-3 months and will involve daily 'play' 'sessions of about 15 minutes. At the end of each play session we will spend approximately 5 minutes reading from enjoyable and educational books when we will concentrate on the learning of new words. All sessions will be videotaped for later analysis. When the study has finished, a summary of the findings will be given to parents whose children take part. As a token of our appreciation of childrens' participation in the research we shall also provide, in consultation with parents, a gift for each child at the conclusion of the study.

As you are aware it is Nursery policy to obtain parental approval for any studies in which the children participate, so we would be grateful if you would take the trouble to complete the slip below and return it to Sue Kennedy (Nursery Manager). Sue will also help to keep you informed about all aspects of the study. You will have the right to withdraw you child from the study at any time. All data will be kept confidential.

Many thanks for your help.
Yours sincerely,

C. Fergus Lowe
Professor of Psychology and Head Of Department

Annwyl Riant /Gwarchodwr,

Yn Nhir na n-Og, rydym ar fin dechrau astudiaeth arall o ddatblygiad iaith mewn plant sy'n adeiladu ar yr astudiaethau blaenorol yr ydym wedi eu cynnal yma. Mae'r ymchwil yn ymwneud yn bennaf â dysgu enwau gwahanol wrthrychau i blant, o fewn cyd-destun 'chwarae'. Mae gennym ddiddordeb mewn dod o hyd i ffyrdd o ddysgu iaith yn effeithiol a gobeithiwn arsylwi sut mae hyn o fudd i ddysgu plant yn gyffredinol.

Gwneir pob ymdrech i sicrhau fod y broses, drwy'r hyn y bydd ein staff yn enwi gwahanol wrthrychau i'r plentyn ac yn asesu ei (d)dealltwriaeth o'r enwau hyn, yn un pleserus iawn i'r plant. Mae'r gweithdrefnau a ddefnyddir wedi cael eu defnyddio mewn amrywiaeth o astudiaethau tebyg ac mae'r plant wedi mwynhau cymryd rhan ynddynt yn fawr iawn. Roedd nifer o'r rhieni yn gadarnhaol iawn am werth y rhyngweithiadau un i un beunyddiol hyn.

Cyfarwyddwr yr astudiaeth benodol hon yw Marleen Adema, aelod uwch-radd o'n tîm ymchwil, sy'n gweithio dan fy ngoruchwyliaeth i. Mae gan Marleen brofiad mewn cynnal astudiaethau babanod yn y Feithrinfa. Bydd yn hapus trafod manylion yr astudiaeth gyda chi ac yn rhoi'r wybodaeth ddiweddaraf i chi ynghylch sut mae'r astudiaeth yn mynd rhagddi.

Mae'n anodd pennu pa mor hir fydd yn ei gymryd i arsylwi yr effeithiau iaith yr ydym yn ymchwilio iddynt, ond disgwyliwn y bydd yr astudiaeth yn para tua 2 - 3 mis a bydd yn cynnwys sesiynau 'chwarae' dyddiol o tua 15 munud. Ar ddiwedd bob sesiwn chwarae, byddwn yn treulio tua 5 munud yn darllen o lyfrau pleserus ac addysgol pryd byddwn yn canolbwyntio ar ddysgu geiriau newydd. Caiff yr holl sesiynau eu recordio ar dâp fideo i'w dadansoddi yn nes ymlaen. Pan fo'r astudiaeth wedi gorffen, rhoddir crynodeb o'r canfyddiadau i rieni'r plant sy'n cymryd rhan. Er mwyn dangos ein gwerthfawrogiad o gyfranogiad y plant yn yr ymchwil, byddwn hefyd yn rhoi anrheg i bob plentyn ar ddiwedd yr astudiaeth, ar ôl ymgynghori â rhieni.

Fel y gwyddoch, mae'n bolisi gan y Feithrinfa i gael caniatâd gan y rheini ar gyfer unrhyw astudiaethau maent eu plant yn cymryd rhan ynddynt, felly byddem yn ddiolchgar pe baech yn llenwi'r bonyn isod a'i ddychwelyd at Sue Kennedy (Rheolwr y Feithrinfa). Bydd Sue hefyd yn helpu i roi gwybod i chi am bob agwedd o'r astudiaeth. Bydd gennych yr hawl i dynnu eich plentyn yn ôl o'r astudiaeth unrhyw adeg. Cedwir pob data yn gyfrinachol.

Diolch yn fawr iawn am eich cymorth.

Yn gywir

C. Fergus Lowe
Athro Seicoleg a Phennaeth yr Adran

Name(s) of child or children: ________________________________

Name(s) of parent(s) / guardian: ______________________________

▪▪

Please tick below if you do not wish your child to participate in the naming study:

NO CONSENT [ ]

▪▪

Please tick below if you want the researcher to talk to you and give more information about this study:

MORE INFORMATION REQUESTED [ ]

▪▪

Please sign below, and give your contact details, if you consent for your child to take part in the naming study.

I CONSENT to my child's participation. I understand that for the duration of the study my child will be engaged in one-to-one interactions with the named researcher, conducted in one of the test rooms in the nursery.

Signed: ________________________________

Date: ________________________________

Telephone Number: ________________________________

Enw(au) y plentyn neu'r plant: ______________________________

Enw(au) y rhiant(rhieni) / gwarcheidwad:______________________________

Ticiwch isod os nad ydych yn dymuno i'ch plentyn gymryd rhan yn yr astudiaeth enwi:

DIM CANIATÂD [ ]

Ticiwch isod os ydych am i'r ymchwilydd siarad â chi a rhoi rhagor o wybodaeth am yr astudiaeth hon:

ANGEN RHAGOR O WYBODAETH [ ]

A fyddech cystal ag arwyddo isod, a rhowch eich manylion cyswllt, os ydych yn caniatáu i'ch plentyn gymryd rhan yn yr astudiaeth enwi.

RWY'N CANIATÁU i'm plentyn gymryd rhan. Rwy'n deall y bydd fy mhlentyn yn cymryd rhan mewn rhyngweithiadau un i un gyda'r ymchwilydd penodedig yn ystod yr astudiaeth, a gynhelir yn un o'r ystafelloedd profi yn y feithrinfa.

Llofnod: ______________________________

Dyddiad: ______________________________

Rhif Ffôn: ______________________________

APPENDIX B: DEBRIEFING LETTER IN ENGLISH AND WELSH

Dear Parent/Guardian,

I would like to thank you for consenting for your child's participation in the language development study, at Tir Na n'Og Nursery, Bangor.

The overall aim of this research program is to find out more about how infants and young children learn the names of different objects. We are interested in finding ways to teach language effectively. This particular study is concerned with naming at the basic and superordinate level. Examples of basic level names are 'dog' and 'cat'. What we call 'dogs' and 'cats' at the basic level, are 'animals' at the superordinate naming level.

On the presented DVD you will find a selection of the recordings featuring your child and illustrating our procedures. The study consisted of three stages. First, the children were taught names for eight 'alien' animals, divided over four categories. Every category was made up of two different animals, like two different types of dogs, still belonging to the same category: dog. The aliens were newly designed and the names were unfamiliar (*Hib, Tor, Feb* and *Lup*). This was to ensure that the children had never encountered either the aliens or the names before. The animals were presented on a carousel, four at a time. We found that 3-4 year old children were able to learn to name the newly designed animals.

The second stage of the study consisted of a word game. In this word game two of the four basic category names (*Hib, Feb, Tor, Lup*) were linked to two different superordinate names. Like dogs and cats are all animals on the superordinate naming level, we used the names *Zaag* and *Noom* to designate the new aliens on the superordinate level. In the word game, the superordinate level names were linked to the basic level names by having the child repeat a verbal chain. For example, *Hib-Zaag-Hib-Zaag* and *Lup-Noom-Lup-Noom*. After several trials the child was prompted to produce the complete verbal chain after only hearing the first name.

The last stage of the study was testing whether this simple word game would make the child able to point at the right animal when asked *Where's the Zaag?* or *Where's the Noom?* The results were encouraging showing that children can then select the correct animal when asked to do so even when the animal had not been given the superordinate name directly.

This study will be replicated with more children in the future. In later studies more research will be done on what are necessary and what are sufficient aspects of the language environment for the child to learn to name animals and objects on both the basic and the superordinate level.

If at any time you wish to learn more about our findings, please do not hesitate to contact me directly (tel. ..., email ...), or through Mrs. Sue Kennedy at the Nursery, to arrange an informal meeting at a time convenient for you. Please accept the small gift as a token of our appreciation. And once again, thank you ever so much for your help.

Yours sincerely,

Marleen Adema

Annwyl Riant/Gwarcheidwad,

Carwn ddiolch i chi am gydsynio i'ch plentyn gymryd rhan yn yr astudiaeth datblygiad iaith ym Meithrinfa Tir Na n'Og, Bangor.

Nod y rhaglen ymchwil yw dod i wybod mwy am sut mae babanod a phlant mân yn dysgu enwau gwahanol bethau. Rydym eisiau darganfod ffyrdd o ddysgu iaith yn effeithiol. Mae'r astudiaeth hon yn ymwneud ag enwi ar y lefel sylfaenol ac uwch-fesur. Esiamplau o enwau ar lefel sylfaenol yw 'ci' a 'chath'. Mae'r hyn rydym ni yn alw yn 'gŵn' a 'chathod' ar y lefel sylfaenol, yn 'anifeiliaid' ar y lefel enwi uwch-fesur.

Ar y DVD, fe welwch ddetholiad o recordiadau o'ch plentyn ac yn dangos ein trefn o weithio. Yr oedd tri cham i'r astudiaeth. Yn gyntaf, roeddem yn dysgu i'r plant enwau wyth anifail 'dieithr', wedi'u rhannu dros bedwar categori. Roedd pob categori wedi ei wneud o ddau anifail gwahanol, fel dau fath gwahanol o gi, yn dal i berthyn i'r un categori: ci. Newydd eu cynllunio yr oedd yr anifeiliaid dieithr, ac yr oedd yr enwau yn anghyfarwydd. (*Hib, Tor, Feb* a *Lup*). Y rheswm am hyn oedd gofalu nad oedd y plant wedi dod ar draws yr anifeiliaid dieithr na'u henwau o'r blaen. Cyflwynwyd yr anifeiliaid ar garwsél, bedwar ar y tro. Cawsom fod plant 3-4 oed yn medru dysgu enwi'r anifeiliaid newydd.

Gêm geiriau oedd ail gam yr astudiaeth. Yn y gêm hon, yr oedd dau o bedwar enw'r categorïau sylfaenol (*Hib, Feb, Tor, Lup*) yn cael eu cysylltu â dau enw uwch-fesur gwahanol. Yn yr un modd ag y mae cŵn a chathod oll yn anifeiliaid ar y lefel uwch-fesur, roeddem yn defnyddio'r enwau *Zaag* a *Noom* i ddynodi'r anifeiliaid newydd dieithr ar lefel uwch-fesur. Yn y gêm geiriau, cysylltwyd yr enwau lefel uwch-fesur i enwau'r lefel sylfaenol trwy gael y plentyn i ail-adrodd cadwyn eiriol. Er enghraifft, *Hib-Zaag-Hib-Zaag* a *Lup-Noom-Lup-Noom*. Wedi llawer o dreialon, anogwyd y plentyn i gynhyrchu'r gadwyn eiriol lawn wedi clywed yr enw cyntaf yn unig.

Cam olaf yr astudiaeth oedd profi i weld a fyddai'r gêm syml hon yn gwneud i'r plentyn bwyntio at yr anifail iawn pan fyddai rhywun yn gofyn *Lle mae'r Zaag?* neu *Lle mae'r Noom?* Yr oedd y canlyniadau yn galonogol, yn dangos y gall plant ddewis yr anifail cywir pan ofynnir iddynt wneud, hyd yn oed pan na roddwyd yr enw uwch-fesur yn uniongyrchol i'r anifail.

Bydd yr astudiaeth yn cael ei hail-adrodd gyda mwy o blant yn y dyfodol. Mewn astudiaethau yn nes ymlaen, bydd mwy o ymchwil yn cael ei wneud ar ba agweddau o'r amgylchedd iaith sy'n angenrheidiol a pha rai sy'n ddigonol i blentyn ddysgu enwi anifeiliaid a phethau ar y lefel sylfaenol a'r lefel uwch-fesur.

Os carech wybod mwy am ein canfyddiadau, cofiwch fod croeso i chi gysylltu â mi yn uniongyrchol (ffon ..., e-bost ...), neu trwy Mrs. Sue Kennedy yn y feithrinfa, i drefnu cyfarfod anffurfiol ar adeg cyfleus i chi. Dyma anrheg fach fel arwydd o'n gwerthfawrogiad. Ac unwaith eto, diolch o galon am eich help.

Yr eiddoch yn gywir,

Marleen Adema

APPENDIX C: CHILDREN'S COMMENTS

Study 1

Cheryl[1]

- (About the aliens:) Where do them live?
- Where does them live?
- E: They live in different places ... they hide.
- Does these talk? E: They make noises.
- Did you make these? E: Yes, I made these.
- C points at two animals: Why don't them have a mouth?
- E: Well, they're all different.
- Can we do something else now, 'cause I'm bored.

Davey

- About Feb 1: "Looks like a doggie. It's got a tail!"
- Points at worksheet with aliens: "You've got pictures of them!"
- About Tor 1: "A fishy!"
- About Hib 1: "It's got big feet!"
- While smiling he calls Tor 1 a feb, then he tacts correctly.
- E about the aliens: "We put them on the carousel".
- D: "What does that mean?" E: "What does what mean?"
- D: "Carousel."
- After one trial block, E: "That was really good, Davey! Do you want to play it again?" D: "No, thank you."

- During tact training, D starts talking about the body parts of the aliens, asking what the animal can do with his horns, and tail, and with the spikes on it's back.
- Later he asks: "What do lups do?"
- E: "Oh, various things. They can walk, and jump, and things like that.
- D (about Feb 2): "What can he do?" E: "I think they do the same."
- "Why have they got tails?"
- E: "Some have tails, not all of them."

- Near the end of a trial block, E: "Only two more ..."
- D: "No, a few more!" (D thinks a few is less than two.)
- When reading a story, E points at a number and asks: "What's this?"
- D (with a big smile): "A hib!"
- (About Feb 2) "Looks like a tree."
- In a trial block with mixed sets and the febs exchanged, D points at Feb 1: "That looks like an old animal." (meaning: an animal from Set 1)

[1] Sessions with Cheryl were discontinued (see Study 1a, participants).

- The experimenter excitedly points at D, saying "yes!" when he keeps giving correct responses. When E does this repeatedly, D says: "Stop pointing at me!"
- "I like hibs."
- D keeps making errors when Hib 1 and Hib 2 are targeted. After another error on
- Hib 1: "I'm going to remember it now!"

Jon

- About the Lup (while pointing at his mouth): "He might talk."
- E points at the Hib and asks what it is.
- J: "I don't know."
- E: "It's a Hib."
  J: "A Hib Hib Hooray!"
  In a later session, when E points at an animal and asks "What is this?",
- J first says "Hib Hib Hib Hooray!", and later also: "Feb Feb Feb Hooray!"
- After 20 sessions, J asks: "Are they toy animals?"
- In the session where the Lups are exchanged, J points at the Lup:
- "That's funny! That one!"
  E: "It's a special day today ..."
  J: "That's why you put that on the other carousel, yeah?"

- When E takes one carousel off the table and puts the other one on, J points at both Lups: "That's the same animals! That's a Lup, and that is a Lup!"
- About Hib 2: "That's a naughty girl". E: "Yes, it's a naughty one."
- J: "And that's a good boy."
- About Tor 2: "Is that swimming in the water?"
- (During maintenance trials tacts Hib 2 correctly, then says:) "Hib-zaag-hib-zaag."
- (J turns over the carousel and sees it's the lid of a biscuit tin:)
  "Why are you pretending this is a carousel? E: Because it works like a carousel."
- "Marleen, when I was on holiday I was thinking about your animals."

- (In gesture maintenance training, Tor 1 targeted, tacts tor at higher level:) "Zaag."
- (In gesture maintenance training, Tor 1 targeted:) "Tors go like this."
- (In animal cry maintenance training, Feb 2 targeted)
  "It's a feb and it says Raagh!"
- (In animal cry maintenance training, Hib 1 targeted)
  "It's a hib and it says Boo!"
- (Plays) "The hib jumped over the feb."
- "The hib is pulling the feb's tail."
- "They're looking at you, not me. They don't like me."

- (During maintenance training for animal cries) "It's a feb, and it does (hand on
- shoulder). E: "And what does it say?" J: "Boo!"
- (Feb 2 targeted) "Feb, Raagh!"
- (Hib 1 targeted) "Hib, Boo!"
- (In maintenance training for animal cries:)
  "The feb goes Raagh!, and that's funny, the hib goes Boo!, that's funny."
- (In maintenance training for animal cries, Feb 2 targeted:)
  "Feb, and it says: Raagh!"
- (In maintenance training for gestures:)
  "Lup, and it goes like this (produces gesture)."
- (In maintenance training for animal cries, about Hib 1:)
  "Boo! … it's too funny for me!"
- (In maintenance training for animal cries, about Feb 2:)
  "It does (hands in front), and it's a feb, and it says Raagh!"

- (In maintenance training for animal cries, Feb 2 targeted)
  "Raagh! … Aah! I nearly scared myself!"
- (Plays) The hib was pulling the feb.
- (At the end of a session we play hide and seek. J says I need to count to 150.
  (At another occasion) I hide underneath J's coat, J seeing my feet:
  "You're too big for little things, Marleen!"
- J switches light off at end of session. E: "Hey, what are you doing?" E switches
  light back on. J: "You might waste electricity."
- (In maintenance gesture training, Tor 1 and Lup 2 presented, Lup 2 targeted.)
  "Feb-noom-feb-noom … lup-noom-lup-noom."
- (E reads story at end of session, and always closes the book by slapping it.
  J: "Marleen, I think you've got too much air in your room."
  E: "I don't think you can ever have too much air in a room."
  J: "You do when you slap your book!"

Wendy

- In session 6 the child says about the animals: "These are not real, no?"
  E: "You know what? They are!"
- E points at an animal and asks "What is this?" Child: "A little Feb."
  And later: "A little Hib."
- "The Hib is making a noise!"
- (About Lup 1) "He nearly bit my finger!"
- (About Tor 2) "That's a naughty tor."
- (About Lup 2) "That one has lots of noses!"
- (In an additional session with only Lup 1 and Feb 1)
  "Where's you tor? Where's your naughty tor?" E: "Oh, he's playing outside."

Lyn

- (In a session with mixed sets, in which the hibs were exchanged, L points at
  Hib 2)"Why have you got one new one?"

Jim

- In the first session with the Febs exchanged, J points at a Feb:
  "Hey, why is that one on here. He lives on the next one over there!"
  He points at the other carousel underneath the table.
- "Where has the other eye gone?"
  E: "He has only one."
- "Where's the Hib-snake?"
- Points at a body part of a Tor: "How does he have a leg there?"
- In maintenance training about Hib 2: "Mr. Hib."
- About Tor 2: "Mr. Tor."

Sebastian

- First reaction to the aliens:
  "Are they dinosaurs? What are they? Are they dinosaurs?"
- In the first session he describes the animals: "It's got a tail / one eye / two feet."
- "Who been making these (animals)?" E: "I made them."
- S: "Why have you been making them?" E: "Because I like making animals."
- S: "Have you got any more?"
- "Have you been making the eyes as well?" E: "Yes."
- S: "How did you make them?" E: "Out of clay."
- "Why have you made them out of clay?" E: "Because I like red clay."
  S: "Why have you made them out of red clay?" E: "Because I like red."
- At introduction of second set, S points at one animal: "That's a long one."
  Points at another one: "Looks like a mousie."
- During training stage for Set 2, there is maintenance training for Set 1. After a block of maintenance training, the boy is eager to continue with Set 2: "Are we going to do other animals? Get them on the table!"

- Points at picture on worksheet: "Is that a Lup?"
- "Why have you not got two eyes on that (alien)?"
- In first sessions with mixed sets (Febs exchanged), points at the feb and says:
- "This is a funny shape … (laughing) you put 'm on the wrong one! It should go on that one. (Then he crawls underneath the table to the other carousel and points at the second feb:) and that should go on the other one!" Later he points at the tor on the table and at the one under the table, and he laughs: "Two tors."
- When tors are exchanged, he points at the tor and says: "It's that one! It's silly!"
- Correctly tacts the tor, then points at tor on worksheet:
  "And it's there, in a box!"
- S has noticed that the experimenter ticks off the targeted animals on the worksheet. When at one point she doesn't tick off, S says: "You've got to do a mark on that (worksheet)!"

- (During intraverbal game.)
  S: "Why did you not bring the animals today?"
  E: "Well, sometimes they're asleep and I don't want to wake them."
  S: "Do they live here?"
  E: "Yes, they live here. They live in a little box. That's where they sleep."
- During maintenance training.
  S (when Hib 1 is targeted): "Looks like a mouse."
  E: "Yes, but it's not."
  S: "Looks like it, though."
  E: "But what is it??"
  S: "Hib."
- S starts singing in rhyme about Hib, Feb, etc.

Belinda

- Tacts two separate aliens, Hib 1 and Feb 1: "Kangaroos."
- When a Tor is targeted during maintenance training: "A Hippopotamus."
- When a hib is targeted In the 24th block of training for the first set: "A Rhinoseros."
- In general about the aliens: "They must have two eyes. That one only has one!"
- E: "They all have one eye." B: "Why?"
- E: "Well, some animals have one eye and some have two."

Ellie

- When told an animal is called a "Hib", the child says: "Looks like a Hippopotamus!"
- In session 33 the child starts calling the animals "Torry", "Febby", "Hippy" and "Luppy".
- In session 34 the child recalls that a month earlier she did not know the names of the animals yet, or as she puts it: "When I was little I didn't know these names, right?"
- When the second set of animals is introduced, and the experimenter says about the first animal "This is a Hib", the child reacts: "Oh, like the other one."
  At the end of that session she says: "I like your new animals."
- As late as in the stage of mixed sets of tact training, Ellie calls Feb 1 "a zebra".

Sara

- After one correct response Sara cheers and says: "I can do it!"
  E: "Of course you can, you're a smart girl."
- "Why has it got one eye?" E: "Because it's got to see, just like you."
- S: "Has it got a nose?"
- E: "Do you want to play a bit more, read a story, or go back to the other children?" S: "Do this!"

- When a feb is targeted, S replies: "hib".
  E: "That's a feb."
  S: "A hib". E: "A feb".
  S: "Pretend it's a hib."
  E: "No, it's a feb. What is it?"
  S: "A hib and a feb."

Alun

- Experimenter: "You've got an animal on your jumper. What is that?"
  A: "I'll show you. It's a dragon."
  E (scared): "Oh, a dragon!"
  A: "If you make one, you can put it on there!" He points at the carousel.
- When the Lup is targeted in session 31, A says:
  "That's a funny little Lup, isn't it?"
- "I'm not going to mess about this time. I've been a good boy for my mummy and now I'm going to be a good boy for you!"
- When the Febs are exchanged: "Did you make the Febs jump over?"
- During the listener probe session, the second experimenter asks:
  "Where's the Feb?"
  A points: "There! He's looking at you."
  Then he adds: "Naughty Feb. He doesn't want to play."
- A tells E to fix the door, because it squeaks, and he explains how it should be done.
- After review trials before a test,
  E: "Ok, let's sort things (the aliens) out for Vicky (E2)."
  A: "Yes, cause Vicky doesn't know how to sort things out."

Kyle

- About the aliens: "Why did you make them of stones? Cause they're so hard!"
  E: "I made them out of clay."
- "Why's it got horns there?" E: "Well, some animals do."
- (About Lup 1) "If it's got a tail, it's a doggie."
- "Next time you make some dinosaurs!"
- When the hib is introduced at the start of training of Set 2: "Oh, a different hib?"
- When mixed arrangements are presented for the first time, points at carousel and says: "Why have you got old ones on there? Why have you got a new one on there?" Then he points at the other set and says: "That other one is a lup, with horns on it!"

Cameron

- About Lup 1: "Looks like a rat."
- In the second session, about the aliens: "Aren't we making a home for them?"
  E: "Well, they have a home. And during the day they want to be on the carousel. They want to spin around.
  C: "But they don't want to have a home?"
  E: "Maybe they do ... but they have a place to sleep."
  C: "Where?"
  E: "That's a little box."
  C: "Where?"
  E shows the box.
- In session 33, C says about the Lup, which he consistently calls "a Lump": "Why don't you put a lump on that one, so I know it's a lump."
- When there's no progress for a while in tact training, the Thomas the Tank Engine music book is introduced. For every correct response a finger is held up. C gets to press one button for every three correct responses. It worked marvellously! After a while C suggests a different reinforcement schedule: Now you have to have five (fingers up for me to press a button).

Sasha

- In the session 3: "They have only one eye. I've got two!
  Aliens have only one eye too."
- Calls the Lup a "Luppy".
- In session 18, S calls the Feb "a giraffe".
- E explains that S needs to concentrate, S: "My brain is not in!"
- After tact test trial block 1, S correctly says: "I had all of them right."
- In a transfer of function test, two aliens are presented.
  S: "One from one carousel and one from the other."
- "Other people don't know the names of these animals, right?
  Only the people who made them."

- At the end of a session, S: "Are you going home now?"
  E: "No, not yet, I'm going to the university first."
  S: "Do you play with the people in the university as well?"

Study 2

Tomos

- About Feb 1: "Looks like a kangaroo."
- In pre-training for the word game. Instead of "vegetable-food", T says: "vegetable-tea".
- In training for the alien name word game: "feb-animal", in maintenance training for the pre-training of the word game: "fish-noom".

Ginny

- About Hib 1: "He goes in the water."
- When presented with Set 2 and told the names, she says: "A different lup!" and "A different hib!"
- About Tor 2: "That's a real tor."
  E1: "Why?"
  Ginny: "Cause he's got a tail, spikes … (describes further)"
- "That's a real feb/hib/lup" – describes features.
- Points at Hib 1 (who has a smiley mouth): "He's happy!"
  Points at Feb 1: " … and he … oh, he hasn't got a mouth!"
- Counts spikes on Tor 2.

- About Lup 2: "He's having a bath." (2x)
- About Tor 2: "He looks like a doggie."
- When it had been quite a while since maintenance training was done for the yellow shapes, G starts using alien names for them.
- At some point, Ginny sings: Feppy feb … hippy hib, torrie tor, luppy lup.
- In training sessions for the word game of Leg 2, when E puts Lup 1 on the table, G says: "He's a noom!"
- During a training session for the word game, maintenance training is done for tacting the aliens. After correctly tacting Lup 2, she says "lup-noom", and she does the same for Tor 2 "tor-noom" and for Feb 2 "feb-zaag".

Huw

- When the tors are exchanged, Huw points to a tor and says: "That's not right!"
- When first presented with the aliens, and told the names, he calls the tor a dragon, the lup a reptile, the hib a bat, and he says that the lup is "like a croc", and the feb is "like an egg".
- Sometimes children joke by giving the wrong reply (and then showing a very big grin while checking the experimenter's reaction). Huw: "fruit-animal" (laughs).
- In function training for Study 2d, Huw: "Raagh … like a tiger".
- Later he starts to make gestures along with the cries; the peek-a-boo gesture for "Boo", and putting his hands up in the shape of tiger claws in the direction of the experimenter. He does this more often, also in a session the day after.

- In no reinforcement trials, when presented with two aliens: "feb and hib".
  Then maintenance training is done, E1: "We're going to play the next game..."
  Huw: "…with the tor and the lup."
- In maintenance training for 2c (during Study 2d) when Lup 1 is targeted:
  "Lups do this?"
- In maintenance training for 2c and 2d (after first test for 2d), in two trials he first tacts the aliens ("tor" in a trial for 2c gestures, "hib" in a trial for 2d cries).

Glyn

- About Vegetable 2: "That's dinner!"
- When Reptile 1 is targeted: "He's looking at my belly!"
- When the aliens are introduced…
  Hib 1 is targeted: "What is a hib?" E1: "That's this one."
- About Lup 1: "Looks like an elephant.
  That's his tail and that's his trunk (points)."
- About Feb 1: "A donkey!"
- About Lup 1: "Why does it have an eye?"
- When Tor 1 is introduced: "What is a tor?"
  E1: "This is a tor!"
  Glyn: "He goes on the sea."

- Glyn calls Feb 1 "fluf".
  E1: "Almost right, it's a feb."
  Glyn: "What is a feb?"
- Later he calls Feb 1 "fluf" again, and says:
  "It's got fluffy bits on it. It's a doggy."
- About Tor 2: "That swims, it's a carrot."
- Calls Feb 1 "febby".

- Glyn: "You're a funny one."
- E1: "Why?"
- Glyn: "Cause you got a funny necklace."
- E1: "Do you like it?"
- Glyn: "Yes, and I like all your clothes!"
- E1 (stunned): "Ah, that's a very sweet thing to say!"

- When reading a story at the end of a session, a monkey in the book says "I've got 10 strong fingers, just like you".
  E to Glyn: "Have you got 10 fingers?" (Many children when asked this question, start counting their fingers.)
  Glyn: "Yes."
  E1: "How do you know?"
  Glyn: "Cause I couldn't eat my dinner if I would have only one!"

  (In reply to the same question, "how do you know?", in the same context, Jon once said: "Cause I count my fingers all the time!")

- At tea time, E says to Simon (who will be leaving the nursery soon):
  "I will miss you when you're gone."
  Glyn (age 3): "I will look after you."
- In maintenance training for tacting the aliens, Glyn calls Tor 1 a hippo, and a carrot.

Kevin

- In pre-training (tact training for the familiar shapes) when the mushroom/vegetable is targeted, Kevin:
  "Vegetable … that's called a dinner! Mushroom!"
- When the crocodile is targeted: "Reptile … called a crocodile. Snap!"
- Mushroom targeted again: "Dinner! … vegetable … I don't like those."
- Later in training, when the mushroom is targeted: "Fruit", then Kevin shakes his head and says, "That's not a fruit, is it?"
  E: "No. What is it?"
  Kevin: "Vegetable … dinner, mushroom. I like it."
- Kevin generalised what he learned during pre-training (the word reptile for a crocodile and snake), to other contexts. When reading a book about a little stripy jeep in the jungle, E points at a crocodile and says, "What is that?" Kevin: "A crocodile", then he adds, with a smile: "… and it's a reptile!"

Leona

- When the aliens are introduced, Leona starts trying for the correct response by using nonsense words like "wub" for Lup 1 (2x) and for Hib 1 (1x), and real words like "chin" for Lup 1 (1x), for Hib 1 (3x) and for Tor 1 (1x), "toy" for Tor 1, "pub" for Lup 1, and later "web" for Hib 1.

Adam

- In tact training for familiar shapes, Adam called the vegetable a reptile (looking at experimenter with a big smile).
- Adam in tact training with the familiar shapes:
  "I don't have any vegetables in my house, only fruits."

- At the end of a session, Adam picks the book about planes at an airport. On the cover of the book there was a picture of a wind sock. Adam pointed at it and said: "That is not in the book." Experimenter: "Yeah, that's a pity." Adam (holding up his hands, while asking firmly): "So where is the pity???!"

- At end of a session with the word game for alien names, E1 says: "You were brilliant!" Because Adam doesn't react, E1: "Did you hear me? You were brilliant! … Adam, did you hear me? What were you?" Adam: "A hero!"

Cecelia

- Calls Tor 1 a tortoise or a tortle, and Hib 1 a broom.

Louise

- When Lup 1 is introduced: "Looks like a dinosaur."
- About Hib 1: "Looks like a huck." Later: "Like a hippopotamus."
- About Feb 1, L starts rhyming: "Like a spider web."
- About Lup 1: "Like a pub."
- About Lup 1: "Lup's got bumps on."
- In maintenance training for alien tacting during pre-training for word game, about Feb 1: "That's a boy!" E1: "Why?" Louise: "Cause it's got a tail."

- E: "When's your birthday, Louise?"
  L: "When I'm 4."
  E: "Oh, right. How long is that? How many nights?"
  L: "6."
  E: "Oh, really? That's good, cause that's quick!"
  L: "And 7, and 8, and 9 … 9 miles!"
  E: "Nine miles? Nine miles till your birthday? "
  L: "Yeah."
  E: "Oh, okay."
  L: "Yeah … and it's miles away … so I have to go on an airplane …"
  E: "To get to your birthday?"
  L: "Yeah."
  E: "Oh, okay. Where's your birthday then?"
  L: "At December."
  E1: "Oh, and you're going with an airplane to December?"
  L: "Yeah, to my birthday. … it was someone's birthday … (thinks) I went to Chrissy's birthday. It was only down the road from my house."

- In the word game with alien names, Louise says "hib-animal" and is corrected by the experimenter telling her it's "hib-zaag".
  Louise: "A hib is also an animal, isn't it?"
- Louise: "Hibs are rocks … in Welsh." And: "Febs are spider webs in Welsh."

Jamie

- In maint training for tacting Set 1 of the familiar shapes, Jamie points at the fish and the reptile and says: "These two are animals."
- Jamie and E hear the other children coming in to the corridor to get ready for lunch. E: "Do you want to read a story, or do another game with these quickly?" Jamie: "Do this!"

- Jamie about Lup 2: "Looks like an alien."
- Later: "Luppy is a naughty one."
- When Feb 2 is targeted: "L for hib … I was joking. Feb."
- Calls the animals "luppy", "torry", and "hibby".

- Jamie looks at one of the camera's: "If you climb into that camera, you'll turn into a picture!"

Mike

- In pre-training of Study 2c (function training), when asked how the crocodile goes, M makes the gesture and says "floating". In the transfer of function test, when asked how the crocodile goes, M makes the gesture and says "it slithers on the ground". Later in the test, when the shark is targeted, M makes gesture, and says "shark, swims". When the snake is targeted: "Does a snake swim?"
- In Session 1 of (alien) function training in Study 2c, M makes the hands in front gesture, and says "bridge!". Then he starts singing "London bridge is falling down…" In Session 3, he again says "bridge" when he makes that gesture.

- After the first block of training trials in Study 2d (function training with cries), maintenance training is conducted for tacting of the aliens, and for the word game. Mike starts integrating all the games in the tact trials. Apart from saying the lower-level name for each of the aliens, he adds for Lup 1 "Boo!", for Hib 1 he produces the LHS gesture, for Feb 2 he says "Raagh!, feb-zaag, Raagh!", and for Tor 2 he produces the hands in front gesture. In all these cases he produces the trained function (either the gesture or the cry). But for Tor 2 he also adds "Boo!" which is an untrained but correct response for Tor 2, so he shows transfer of function in this single trial. In later sessions, he shows more examples of integration of games.
- He also calls the aliens "hippy", "torry", and "feppy".

Simon

- In Session 1 with the familiar stimuli, when prompted "what's this?", Simon frantically starts describing features of the targeted stimulus.
  For the fish: "It goes in the water." For the mushroom: "It grows in the ground and you can eat it … etc.
- S: "I like playing with you!" E: "I like playing with you too!"
- Simon calls Feb 1 "fib".
- When Hib 2 is targeted: "Starts with h…"
- Simon about Tor 2: "Big laughing tor."
- When Hib 2 is targeted: "Big hib." Then Simon points at the picture on the experimenter's worksheet and says: "And a little hib."
- In a trial in which Lup 2 is targeted, Simon checks out all animals one by one, then says "lup" (as if he's naming by exclusion). He does that several times.
- When tors are exchanged: "The tor is a different one!"
- About Feb 1 (the long shape): "Looks like a carrot." Later in the experimental procedure: "It looks like a carrot, but it really is a feb."

- After the fruit-food link in pre-training (word game with familiar names), the link fish-animal is trained. (Some children get their tongue tied with the fruit-food link, as did Simon.)
  S: "I like the fish one better than the fruit. I've got a fish hat on today. Is that why you decided to do the fish one today?"
- During the word game with familiar names,
  Simon: "Is it fish-animal, because a fish is an animal? And it's fruit-food, because fruit is a food?"
  E: "Very clever."
  Simon: "I was thinking that last night."

- During test within with familiar stimuli, Simon picks up the banana, and makes it fly: "I'm making it into a bug."
- Simon touches the snake: "I like the pattern."
- And about the fish: "I like how the clay feels."
- About Lup 1: "It has a tail like a dragon", and about the same alien: "the lup is the alien's dog."
- In the word game, when asked "what do you say, when I say 'hib'?",
  Simon: "I don't remember, I need a clue."

- In the alien name word game: "Why do you say lup-zaag-lup-zaag?"
  E: "Well, that's the game."
- In the word game with the alien names, Simon: "What does 'zaag' mean?"
  E: "Oh, it's just a name."
  S: "Why do they look different?"
  E: "Well, they just do."
  S: "But why do they have the same name?"
  E explains that not all dogs look the same, but they do have the same name.
- In that same session, in maintenance tact training for the aliens, Simon correctly tacts Lup 1 and then says: "It's a zaag, zaag is the second name for lup."
  When Hib 1 is targeted in maintenance tact training,
  Simon: "hib, … noom in the word game."
- When Lup 2 is targeted in maintenance tact training,
  Simon: "lup's second name is zaag."

- During maintenance tact training, about Tor 2: "The tor looks like he's smiling."
- In Study 2c, in function training with the aliens, S is told how the feb goes.
  S: "How does the tor go?"
  Later in that trial block, about how the feb goes (hands in front):
  "It's a bit like the snake one." In training block 3, about the feb and the 'hands in front' gesture: "How does that one do that?"
- In Study 2d, function training with animal cries, Simon points at the animals and says "That one is scary, and that one does peekaboo". (naming the gestures)
- In review trials before the test for Study 2d, trials are done for listener behaviour at the higher name level (which one is the zaag/noom?). Simon whispers the word game when he chooses an alien. At the end of one of the trials when taking away the two aliens, E asks, in reaction to his selection of a stimulus, "That one?" Simon: "Yes, it's the feb".

- In Trial 1 of the transfer of function test for Study 2d, Simon is presented with Tor 1 and Feb 2 while function training for him had involved Lup 1 and Hib 2. Simon: "I need a clue." E: "Well, it can either say Boo! or Raagh!"
- When presented with Pair 2 of the aliens (Lup 2 and Tor 2) in Trial 1 of sheet 3, Simon says again: "I need a clue." E1: "Look at the animal, think of all the games, that can help you. And just do what you think."
  Probably because he doesn't get any feedback on his performance, Simon says at some point during the test: "It doesn't really matter what I say, does it?" E1: "It does, Simon, please think carefully!"

Lee

- When first presented with the aliens, Lee: "They look like monsters."
- In early training, he calls feb 1 regularly a 'fib'.

- In the first sessions with mixed sets, hibs exchanged, in the trial targeting Hib 1, Lee: "Hey, there's a new one (from Set 2) on there! How did it get there?"
  E: "It jumped on!"
  Lee: "No, it didn't. They're not magic!"

- Halfway a pre-training word game session Lee kisses the table. E1: "Hey, did you kiss the table?" Lee: "Yes." E1: "Why?" Lee: "Cause I'm in love with it!"