Development of metaphor and metonymy comprehension: Receptive vocabulary and conceptual knowledge

Gabriella Rundblad and Dagmara Annaz

1Department of Education and Professional Studies, King's College London, UK
2Centre for Brain and Cognitive Development, Birkbeck University of London, UK

Figurative language, such as metaphor and metonymy are common in our daily communication. This is one of the first studies to investigate metaphor and metonymy comprehension using a developmental approach. Forty-five typically developing individuals participated in a metaphor–metonymy verbal comprehension task incorporating 20 short picture-stories. Cross-sectional trajectory analyses linking task performance to either chronological age or receptive vocabulary (mental age, MA) were used to compare the development of metaphor and metonymy. Results showed that development of metaphor and metonymy comprehension is strongly linked with chronological and MA, but metaphor comprehension develops at a slower rate compared to metonymy. It was also found that participants, across all ages, consistently showed around 21% better performance on metonymy. The relationship between metaphor and metonymy comprehension is discussed in terms of linguistic and cognitive models of figurative language comprehension arguing that metonymy is cognitively more basic than metaphor.

Figurative language is a ubiquitous part of everyday communication and everyday life, and is therefore readily used and understood by ordinary speakers and listeners (e.g., Pollio, Barlow, Fine, & Pollio, 1977). Yet figurative expressions have commonly been treated as deviations from some literal equivalent (e.g., Corbett, 1990), with metaphors referred to as manipulations of meaning that children need to grasp in order to enjoy creative literature and poetry in particular (e.g., Smith, 1976). Figurative language studies have also persistently focused on a narrow range of figurative language devices, mainly metaphors, resulting in an incomplete and potentially skewed picture of figurative language comprehension and production. In particular, development of metonymy comprehension and production has fundamentally been ignored.

*Correspondence should be addressed to Dr Gabriella Rundblad, Department of Education and Professional Studies, King’s College London, Waterloo Road, London SE1 9NH, UK (e-mail: gabriella.rundblad@kcl.ac.uk).

DOI: 10.1348/026151009X454373
In the current study, we emphasize a developmental approach by constructing cross-sectional trajectories of children as young as 5 years of age to 37-year-old adults to examine development of metaphor and metonymy comprehension. To this end, we will employ a methodology that highlights developmental change over time (Karmiloff-Smith, Thomas, Annaz, Humphreys, Ewing, & Brace, 2004; Thomas et al., 2009). Our study is of particular relevance due to the limited scope of previous studies that either equate metaphor with figure of speech, or, as in the case of MacKay and Shaw (2004), exclude metaphors, thus limiting the comparability of the study. This is, to the best of our knowledge, the first study to compare and contrast the developmental profiles of metaphor, and metonymy comprehension. Our data reveal for the first time that comprehension of metaphors and metonyms steadily develops into adulthood. Perhaps the most striking result of the current study is the consistent difference in individual performance on the task with metonym scores being around 21% higher than metaphor performance scores.

In the following sections, we will briefly review research on figurative language processing, production, and comprehension, and then consider the development of this ability in typically developing children and adults. We will first, however, need to establish the definition and potential relation between metaphors and metonyms.

**Metaphor and metonymy**

Traditionally, metaphor and metonymy used to be distinguished as relying on similarity versus contiguity (e.g., Ullmann, 1967). We would thus distinguish between *face* in ‘There is a new face on the clock’ versus ‘There is a new face in the office’; here, the former is a metaphor where a part of the body is compared to a part of a clock, whereas the latter is a metonym where the expression *face* stands for the whole person as opposed to just a part of the body (i.e., the contiguity is situated in the spatial part-whole relations of the human body).

Recently, the distinction of conceptual similarity versus conceptual contiguity was reintroduced in combination with prototype theory. Building on Barcelona (2002, 2003) and Dirven (2002), Peirsman and Geeraerts (2006) suggest that metonymy is a prototypically structured concept with the prototypical core of contiguity being formed by spatial part-whole relations (see also Blank, 2003; Feyaerts, 1999; Seto, 1999). Similarly for metaphor, Shen (1992) argues that the comparison involved in metaphorical expressions constructs an *ad hoc* category where one of the terms (usually the second one [e.g., *lion* in ‘Peter is a lion’]) is the prototypical member of that category.

In contrast, cognitive models of figurative language, such as conceptual metaphor theory (CMT), treat figurative language as part of the norm in human communication, arguing that the human conceptual system is figurative in nature (Lakoff & Johnson, 1980). In CMT, metaphor and metonymy are commonly contrasted as a mapping across two conceptual domains versus a mapping within a single conceptual domain (Kövecses, 2002; Lakoff, 1987). Thus, the metonym *face* is based within the domain HUMAN BEING, whereas the metaphor *face* crosses from this animate domain into the inanimate domain CLOCK. The domains are structured by an idealized cognitive model (ICM), and by shifting salience from one concept to another (across domains or within the same domain), particular properties and aspects of the domain(s) and the ICM can be highlighted (Croft, 2002). Unfortunately, domains as well as ICM are quite vague concepts (see Croft & Cruse, 2004; Feyaerts, 1999; Riemer, 2001, for discussions), leaving the metaphor-metonymy distinction somewhat imprecise.
Whether employing a CMT approach, a prototype/categorization approach or a traditional contiguity/similarity distinction, one difference does seem to remain constant; metaphor is more complex than metonymy (i.e., metaphor involves two rather than one domain, metaphor requires the construction of an ad hoc category, or similarity between two concepts is hard to assess unless contiguous relations within each of the two concepts have been determined). We will argue that, from a developmental point of view, it is therefore highly likely that metonymy understanding will exceed metaphor comprehension at any point in time throughout childhood to adulthood. In addition, it is also likely that metonymy develops earlier and faster.

**Figurative language processing**

Early pragmatic studies postulated an initial literal interpretation that only in the event of interpretation failure would trigger a subsequent search for a figurative interpretation (Grice, 1975; Searle, 1979). More recent figurative language processing studies have, however, shown that figurative phrases do not require additional processing time compared to literal phrases (e.g., McElree & Nordlie, 1999). Gibbs (1984, 1990, 2002) and Turner and Katz (1997) have proposed that the processing of figurative expressions is direct and requires neither preceding nor subsequent literal interpretation; a literal interpretation can be forced, but only upon completion of the figurative interpretation. Recent psycholinguistic parallel models (e.g., Frisson & Pickering, 1999, 2001; Glucksberg, 1991; Keysar, 1989; Keysar, Shen, Glucksberg, & Horton, 2000) suggest that both a literal and a figurative interpretation are available and processed at the same time, and that selection is a matter of familiarity and salience. Frisson and Pickering (2007) have shown that familiar metonyms are easier to process than unfamiliar ones, though context can render an unfamiliar metonym familiar. Similarly, interpretations of literal versus metaphorical language show no difference in brain activation if ample and relevant context is supplied (Coulson & Van Petten, 2002; Pynte, Besson, Robichon, & Poli, 1996). The role of salience has been elaborated upon using Giora's (1997, 2002) graded salience hypothesis, with results showing that more salient meanings (literal or figurative) were activated earlier and accessed faster than less salient ones (which also can be literal or figurative).

**Figurative language production**

Early studies of figurative language predominantly focused on the production of similes (e.g., 'It's like a snake' of a winding river, Piaget, 1962) and metaphors (e.g., 'I'm a big waterfall' when sliding down, Carlson & Anisfeld, 1969). At first, data tended to be extracted from recordings of spontaneous speech. In later studies, various elicitation methods were used. It should be noted that many studies are confounded by interchangeable use of metaphors with similes, idiom, and proverbs, treating and labelling them *metaphors*. This incongruous use of the term *metaphor* has probably aided the disagreement as to when figurative language production actually arises, with the earliest studies generally concluding that metaphor production emerges as early as around 18–24 months (e.g., Billow, 1981; Carlson & Anisfeld, 1969; Piaget, 1962; Winner, 1979). Whether these so-called 'child metaphors' are truly metaphorical has been hotly debated in light of metaphor and categorization theory advancements (see Gentner (1988) and Vosniadou (1987), for overviews) and some of the studies have critically tried to distinguish metaphors from overextensions (e.g., Winner, 1979).
In general, production studies have struggled to confirm that an utterance is a conscious violation of reality motivated by the perception of a similarity between the two concepts being compared.

**Figurative language comprehension**

Similarly to production studies, early comprehension studies presumed that evidence of analogical linking of concepts based on perceptual similarity indicated presence of metaphorical understanding. Again, use of different methods and classifications brings inconsistencies and lack of agreement in findings. For example, Winner and Engel (1980) used a matching task where classification of an item in terms of perceptual similarity to another item was scored as metaphorical, Vosniadou and Ortony (1983) asked participants to complete similarity statements involving similes, and Billow (1975) used a sentence explanation task with supporting visual stimuli that mixed perceptual and functional metaphors with proverbs. Moreover, the expressions tested were often constructed idiosyncratic metaphorical expressions rather than actual natural language expressions.

As explored by Waggoner, Messe, and Palermo (1985), recall and construction of meaning for both literal and metaphorical expressions is dependent on supporting contexts, preferably extended and appropriately structured stories. Consequently, results from studies such as Billow (1975; see also Winner and colleagues (1976)) may be confounded. Smith (1976) contrasted metaphor comprehension in children aged 11–12 years versus 13–14 years by supplying the children with a metaphor in context and asking them to explain the context and the target metaphor. Unfortunately, Smith’s contexts generally included other figurative expressions in addition to the target metaphors. Since task complexity has a marked effect on metaphor comprehension (Vosniadou, Ortony, Reynolds, & Wilson, 1984), Smith’s conclusion (1976, pp. 238) that metaphor is ‘a new cognitive domain’ for 11- to 12-year-olds is highly questionable.

Using short stories with familiar concrete nominal metaphors as targets, Jones and Stone (1989) found that participants performed better when asked to paraphrase rather than explain the meaning of the metaphor. They notably used prompts in order to force participants to disambiguate responses. One drawback with paraphrase studies is, however, that paraphrasing requires meta-cognitive abilities that generally develop late in childhood and thus the cognitive demands of such tasks could prevent participants from scoring well on the task whatever their level of metaphorical comprehension ability (Vosniadou et al., 1984).

Due to the theoretical, methodological, and linguistic inconsistencies, studies have not surprisingly yielded inconsistent data regarding the age at which comprehension emerges. Reanalysis of results taking these inconsistencies into account would suggest a three-stage development of metaphor comprehension: exclusively literal interpretations between 3 and 6 years old, onset of metaphor understanding between 7 and 10 years old, and an integrated understanding coupled with the ability to verbalize their understanding from the age of 11 years old. Özçaliskan (2005) used story contexts in a forced-choice task to test this. She found that age 4 marked the onset of metaphor comprehension and age 5 the onset of abstract verbal reasoning about metaphorical mappings. These new findings can, to a large extent, be attributed to the test materials used. In particular, target words were all metaphorical extensions of verbs based on the typical metaphorical mappings described in CMT (e.g., Lakoff & Johnson, 1980) elicited from natural language.
The current study aims to examine the dynamics of developmental profiles of metaphor and metonymy comprehension rather than taking a static view of figurative language comprehension, thus developmental trajectories or growth models will be applied (Annaz, Karmiloff-Smith, & Thomas, 2008; Jarrold & Brock, 2004; Karmiloff-Smith et al., 2004; Rice, 2004). Although one would ideally use a longitudinal design to study development, cross-sectional studies can give an initial approximation of respective developmental trajectories (Thomas et al., 2009), and these can then be validated by a longitudinal follow-up. In order to use this design, it is important that the task is sensitive to developmental change (see Annaz et al., 2008). The task designed for this study is sensitive across a wide age range, thus ideally suited to this approach. One of the main benefits of this approach is the ability to establish the developmental relations between different experimental tasks/variables, assessing the extent to which performance on one task/variable predicts performance on another task/variable across development.

**Method**

**Participants**

Forty-five typically developing children and adults (23 female, 22 male; age range 5; 3–37; 1 years old; $M = 17; 1$) participated in the current study. Details can be found in Table 1. Child participants were randomly recruited from mainstream schools in South-East England, and adults from universities and companies in South-East England, by asking the schools, universities, and companies to forward information sheets and consent forms. All participants had English as their first language, had normal or corrected-to-normal vision and hearing, and had no language-learning disability.

**Table 1.** Participants' descriptives and metaphor–metonymy group scores

<table>
<thead>
<tr>
<th>Statistics</th>
<th>CA (N = 45)</th>
<th>Metaphor</th>
<th>Metonymy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>215</td>
<td>4.8</td>
<td>7.0</td>
</tr>
<tr>
<td>SD</td>
<td>116</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Minimum/maximum</td>
<td>63–445</td>
<td>1–10</td>
<td>3–10</td>
</tr>
</tbody>
</table>

Note. CA, chronological age.

The experimental protocol was approved by King's College London and Birkbeck, University of London, College Research Ethics Committees prior to recruitment of participants. Informed consent (including parental consent where appropriate) was obtained before participation and participants were instructed that they were free to withdraw at any time for any reason. No financial compensation was offered.

**Background measures**

The British Picture Vocabulary Scale II (BPVS) is a receptive vocabulary test used from age 3 to 18 (Dunn, Dunn, Whetton, & Burley, 1997). There are a total of 168 trials where each trial contains one target word and four pictures. Depending on the child's chronological age (CA), a subset of the 168 trials is administered (ceiling level 18 years old). In each trial, the child is requested to point to the picture that illustrates the meaning of the target word. In order to confirm that our participants fall within the verbal IQ range expected from a typically developing population, all participants up to the age of 18 years old were tested on the BPVS. The BPVS was administered in a separate session (Table 2).
Table 2. BPVS test age equivalent and metaphor-metonymy subgroup scores

<table>
<thead>
<tr>
<th>Statistics</th>
<th>CA (N = 24)</th>
<th>BPVS (months)</th>
<th>Metaphor</th>
<th>Metonymy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>119</td>
<td>126</td>
<td>3.1</td>
<td>5.6</td>
</tr>
<tr>
<td>SD</td>
<td>36</td>
<td>46</td>
<td>0.95</td>
<td>2</td>
</tr>
<tr>
<td>Minimum/maximum</td>
<td>63–185</td>
<td>66–206</td>
<td>1–5</td>
<td>3–10</td>
</tr>
</tbody>
</table>

Note. CA, chronological age; BPVS, British Picture Vocabulary Scale.

Material and procedure
Comprehension of metaphor and metonymy was tested using a new experimental task. Thirteen metaphor and 13 metonymy stories were constructed. All target words were lexicalized, nominal expressions and were matched on frequency and age of acquisition using the British National Corpus (frequency counts supplied by www.kilgarriff.co.uk) and the MRC Psycholinguistic Database (www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm), with the exception of the metonym Robbie Williams which for obvious reasons does not feature in these compilations.¹ The 26 stories were assessed for comprehensibility through a pilot study with five children aged 5–9 years old. Thereafter, 10 target stories of each type were selected. A familiarity rating of the metaphoric and metonymic meanings of the final 20 target words (e.g., How familiar are you with FLOOD as in There was a flood (i.e., lots of people) outside the shop) using a five-point Likert scale (1 = very unfamiliar and 5 = very familiar) in a sample of 140 UK undergraduate native speakers of English showed no difference in familiarity between the metaphors and metonyms (metaphors: mode = 5, M = 4.45, SD = 1.01; metonyms: mode = 5, M = 4.38, SD = 1.11; Mann-Whitney U = 959, 397.5, p = .232 [two-tailed]).

The final test material consisted of 10 metaphors and 10 metonyms incorporated into 20 short, simple picture-stories (see sample stories in Appendix A). Participants were tested individually in a quiet room, with participants taking between 20 and 30 min to complete the task. Young child participants were offered a break half-way through. Each story was read, section by section, by the experimenter, while presenting the child with one picture per section. The last part of each narrative contains the target metaphor/metonym accompanied by a picture of one of the characters and the participant is asked to report what that character sees, i.e., to describe the scene, using the question: ‘What does Character X see?’. In order to ensure that the participant understood what they had to do at the end of each story and to implicitly instruct the participants that they needed to respond more elaborately than just repeating the target word, a warm up story containing the hyponym pet was used. In this story one of the characters gets a pet, and consequently, the participants are forced to decide what pet stands for (i.e., what kind of pet it is) when describing the scene. In addition, a total of five story-specific prompts were constructed (examples are given in Appendix B). In the event that the participant did not respond, three prompts (A1–A3) were available, whereas if the participant repeated the target word or failed to disambiguate the target word, two prompts could be used (B1–B2). For some stories, it could be argued that the final prompt available (see prompts A3 and B2 in Appendix 2) is too revealing as

¹ Robbie Williams is probably Britain’s most famous pop singer (e.g., he was ranked as the 9th most famous person in the world by British under 10-year-olds in 2005 according to Luton First (www.lutonfirst.com)).
the relative pronoun used clearly relates to the intended figurative meaning (e.g., the pronoun *who* in 'Who is the flood' where *flood* means 'a lot of people'). However, our participants' responses show that participants who interpreted the target word literally usually did not reflect on the pronoun. Those participants who did notice the pronoun commented on it, yet were persistent in their literal interpretation.

Each narrative was constructed such that the figurative meaning of the target word was more salient, whereas the literal meaning was possible but highly unlikely. Thus in the sample stories, Robbie Williams could possibly be visiting and the river could potentially have flooded the museum; however, it is far more likely that Kate found her Robbie Williams CD in the lounge and that people were queuing outside the museum.

The narrative for each story is short: the average number of clauses per story is 14.45 (SD: 2.56) and the average number of words per clause is 6.48 (SD: 0.99). Each narrative is accompanied by three to four simple, hand drawn and coloured pictures, allowing each story to be divided into three to four sections in order to limit memory demands and aid comprehension of the story. The order of the stories was pseudo-randomized avoiding more than two stories with the same type of figurative language in a row, and the order of presentation was the same for all participants.

**Scoring**

For the purposes of quantitative analysis, responses were classified as either literal or figurative interpretations of the target word. Only in those cases where the response clearly demonstrated comprehension of the intended figurative meaning for the target word did the participant score a pass and was classified score 1. In cases where the participant did not respond at all, repeated the target word or gave an idiosyncratic answer, the story-specific prompts were used forcing the participant to supply a figurative or literal interpretation of the story end; thus making sure that the participant could not fail the test due to lack of comprehension. Literal interpretations of the target expression were scored 0.

In order to ensure reliability in coding of the participants' responses, each testing session was videotaped for subsequent scoring and analysis. Two independent raters coded all testing sessions, with an inter-rater reliability score of .917 (Cohen's kappa; *p* < .001) indicating an outstanding level of coding stability.

**Results**

Participants could achieve a maximum score of 10 correct on metaphor and metonymy respectively. Figure 1 depicts the developmental trajectories linking performance with CA for the participants on metaphor and metonym comprehension, respectively. Analyses of covariance within the general linear model were used in the current data study. Cook's distance function was used to identify potential outliers that could influence regression. There were no outliers in the current data set. The data were analysed as follows: (i) we first examined metaphor and metonymy developmental trajectories separately. Each developmental trajectory was modelled by a linear function relating individuals' accuracy scores to CA; (ii) then both trajectories were directly compared in terms of onset and rate of development; (iii) finally, metaphor-metonymy performance of younger participants was plotted against mental ages (MAs) from the BPVS to explore whether performance is in line with a given standardized measure (see Thomas *et al.* (2009), for a similar approach).
Developmental trajectories of metaphor and metonymy: CA
Performance scores were calculated for each participant and the percentage values on metaphor and metonymy were 48 and 70%, respectively. There was no floor effect on either of the tests but around 13% of individuals achieved highest scores on the metonymy task. Details of group characteristics are presented in Table 1. Initial inspection of the data revealed that participants produced reliable linear trajectories with large $R^2$ values (metaphors: $R^2 = .69$; metonyms: $R^2 = .57$), thus explaining a significant proportion of the model variance. Linear trajectories revealed that the participants' accuracy performance on metaphors and metonyms improved reliably with increasing CA (metaphors: $F(1,44) = 95.29$, $p < .001$, $\eta^2_p = .689$; metonyms: $F(1,44) = 58.04$, $p < .001$, $\eta^2_p = .574$).

Comparison of metaphor and metonymy trajectories: CA
A comparison of metaphor and metonymy developmental trajectories indicated that overall performance significantly improved with CA (main effect of age: $F(1,43) = 119.02$, $p < .001$, $\eta^2_p = .735$). Next, we compared the trajectories for performance at onset (i.e., at the earliest age participants were tested, 63 months) and the rate of increase in performance relative to CA. There was a significant main effect of task ($F(1,43) = 24.99$, $p = .001$, $\eta^2_p = .368$) indicating a reliably higher performance on metonymy by youngest participants tested. However, there was no interaction between age and task ($F(1,43) = .42$, $p = .519$, $\eta^2_p = .010$) suggesting that performance on metaphors and metonyms increases at similar rate of development with increased CA. The relative difference between data points on metaphor and metonymy by individual participants was examined and it was shown that individuals performed on average 21% higher on metonyms in comparison to metaphors.

In order to be able to compare metaphor-metonymy performance with CA and MA, we generated a subsection of the full trajectories that would be overlapping CAs with MA ranges (see Figure 2).
Analyses of subtrajectories (63-185 months old) revealed that participants produced reliable linear trajectories with relatively large $R^2$ values (metaphors: $R^2 = .44$; metonyms: $R^2 = .46$). Similar to analysis of full metaphor and metonymy trajectories, overall performance significantly improved with CA (main effect of age: $F(1, 23) = 30.94, p < .001, \eta^2_p = .584$). However, in contrast to full-trajectories results, there was no significant main effect of task ($F(1, 22) = 0.001, p = .970, \eta^2_p = .001$) indicating similar onset of performance, and there was an interaction between age and task ($F(1, 23) = 5.87, p = .024, \eta^2_p = .211$) suggesting that performance on metaphors improves at significantly slower rate than metonyms in relation to increasing CA. These differences in results are most likely influenced by ceiling scores (13%) on metonymy, and produced artefacts in trajectories that mask real difference in parameters such as onset or rate/growth over time.

**Developmental trajectories of metaphors and metonyms: BPVS**

We then examined whether there was a reliable relationship between metaphor-metonymy trajectories and MA measures based on the BPVS language test to see whether the behavioural scores are in line with the developmental state of other aspects of the cognitive system. Due to limited test age for BPVS (upper limit is 18 years old), we analysed the developmental trajectory of 24 individuals (age range: 63-185 months old). As expected, in the typically developing population (since MA is in line with CA: $t = -2.28, p = .033$), scores increased systematically with verbal mental age (VMA) from the BPVS task. Initial inspection of the data revealed that participants produced reliable linear trajectories with relatively large $R^2$ values (metaphors: $R^2 = .30$; metonyms: $R^2 = .39$), albeit lower values in comparison to subscale-CA trajectories. Linear trajectories revealed that the participants' accuracy on both metaphors and metonyms improved reliably with increasing MA (metaphors: $F(1, 23) = 12.13, p = .002, \eta^2_p = .355$; metonyms: $F(1, 23) = 14.43, p < .001, \eta^2_p = .396$). A comparison of metaphor and metonymy developmental trajectories indicated that overall performance significantly improved with VMA (main effect of age: $F(1, 23) = 19.45, p < .001, \eta^2_p = .469$). Similarly to CA subtrajectories, there was no main effect of task ($F(1, 23) = 0.716, p = .407, \eta^2_p = .032$) indicating no difference in performance between metaphor and metonymy at onset but a marginally significant interaction between age and task ($F(1, 23) = 4.11, p = .055, \eta^2_p = .157$) suggesting that rate of performance on metaphors is slower in comparison to metonyms when relative to VMA.

![Figure 2. Developmental subtrajectories (expressed in accuracy scores) on metaphor and metonymy stories plotted according to (a) CA of the participants (months) and (b) BPVS MA.](image-url)
Discussion

Due to theoretical, methodological, and linguistic inconsistencies, early studies have yielded conflicting data regarding the age at which metaphor comprehension emerges. In addition, studies have tended to limit their focus to a narrow range of figurative language devices, leaving common devices such as metonymy largely unexplored. CMT as well as traditional and recent similarity versus contiguity accounts of metaphor and metonymy seem to agree that metonymy is cognitively more basic than metaphor (cf. Lakoff, 1987; Peirsman & Geeraerts, 2006). The present study aimed to compare and contrast the developmental profiles of metaphor and metonymy comprehension, hypothesizing significantly better performance, and potentially also earlier onset and faster rate of development, for metonymy.

Performance on a new metaphor-metonymy verbal comprehension task incorporating 20 short picture-stories revealed a developmental change on both metaphor and metonymy comprehension with CA. For younger participants, trajectories linking comprehension with receptive vocabulary ability similarly showed an increase in performance with increasing VMA. Although performance variability was slightly larger for CA, both CA and receptive vocabulary ability proved highly reliable predictors of metaphor and metonymy performance.

These results confirm for the first time that comprehension of metaphor and metonymy steadily improves throughout childhood to adulthood. They also validate the task as a sensitive measure of developing language abilities in a typically developing population. As we predicted, analyses divulged a consistent difference in individual performance on the task, confirming a difficulty effect with metonymy comprehension being superior to metaphor comprehension by approximately 21%. The advantage of metonymy over metaphor is consistent with a previous study that reported faster processing times for metonymy (Klepousniotou & Baum, 2007).

Our results are in line with the 'greater cognitive distance' between concepts in metaphor (Pauwels, 1999, pp. 256) that seems to be the defining difference between metaphor and metonymy. Similarly, Coulson and Van Petten (2005) suggest that metaphor processing difficulty is associated with a varying degree of cognitive constraints due to the complexity of the mappings across domains and the conceptual integration that needs to take place. It could be argued that superiority on metonymy is an artefact of lexicalization. However, our task contained both lexicalized metaphors and metonyms matched on frequency and age of acquisition.

Categorization studies have shown that hierarchical classifications, such as part-whole emerge very early on (e.g., Rosch, Mervis, Gay, Boyes-Braem, & Johnson, 1976). Combined with metonymy being easier and cognitively more basic, it seems plausible to suggest that developmental onset of metonymy comprehension precedes onset for metaphor comprehension. We tested this assumption by comparing the trajectories for performance of our youngest participants (i.e., 63 months). The results showed a significantly better performance for metonymy at onset (i.e., the level of performance at the point at which we began measurement), which could be interpreted to mean that children understand metonyms earlier than metaphors. We investigated this further by generating a subsection of the full developmental trajectories for participants up to the age of 15 years old, but interestingly, this effect disappeared. In order to explain this ostensible discrepancy, we first need to look at the rate of development for metaphor and metonymy comprehension.
In contrast to full trajectories where ceiling scores on metonymy distorted results, subtrajectories for younger participants revealed that metonymy understanding improves at a significantly faster rate than metaphor understanding, in relation to both increasing CA and increasing MA. The faster rate of metonymy development paired with its ceiling scores could also explain the discrepancy for onset. In our subsample, metonymy developed so much faster already at the point of onset that we found ceiling scores from around the age of 12 years old. It is likely these ceiling scores 'cap' the increasing difference in comprehension scores and thus 'tip' the full trajectory to the extent that a main effect for task at onset is measured.

The current results suggest that metaphor and metonymy start to develop at similar CA. However, the current task was designed for school-age children who already perform successfully on metonymy (minimum: 3 of 10). Thus, it remains inconclusive whether onset for these two figurative devices overlaps. To clarify this, we would need to devise a test that is suitable for preschool children.

While there may or may not be a difference in onset, there is clearly a difference in rate of development. This difference seems indicative of different cognitive, neural and/or environmental requirements for metaphor and metonymy. One of the most striking features of metonymy is its conformity to a limited number of cognitive patterns that each metonym traditionally subscribes to (Lakoff, 1987; Lakoff & Johnson, 1980). Frisson and Pickering (2007) claim that these patterns (or rules) allow speakers to construct novel metonymic senses without difficulty. Interestingly, most subtypes of referential metonymies are attested almost universally (Bradar-Szabó & Brdar, 2003). We would, therefore, expect that once speakers have acquired and learnt to apply metonymic patterns, performance rates will increase rapidly and speakers will at a conspicuously early point be able to achieve very high performance on typical metonyms. The neural basis for the metonymic patterns remains to be explored, but given its role in semantic language comprehension particularly on a sentence level (Kircher, Leube, Erb, Grodd, & Rapp, 2007), the anterior part of the left inferior frontal gyrus will most likely play a prominent role.

Similarly to metonymy, Lakoff (1987) has established recurrent cognitive patterns for metaphor, though these are less universal than those established for metonymy. Metaphor processing has been widely studied in event-related potential and functional magnetic resonance studies. Recently, Stringaris, Medford, Giampietro, Brammer, and David (2007, pp. 150) found activation of the left thalamus that could indicate specific processing needs for metaphor due to its 'ad hoc concept construction and open-endedness'. In sum, posing specific neural processing needs that require a greater degree of environmental input, metaphor comprehension is likely to develop at a slower rate, and potentially even later than metonymy.

Figurative language comprehension is of clinical relevance in developmental disorders and neuropsychiatry (e.g., autism: Dennis, Lazenby, & Lockyer, 2001; schizophrenia: de Bonis, Epelbaum, DeVez, & Feline, 1997). Moreover, in order to further our understanding of the cognitive constraints in metaphor and metonymy comprehension in the typically developing population, it is crucial to use cross-syndrome comparisons (Karmiloff-Smith, 1998). Preliminary analysis of a related study on figurative language comprehension in Williams Syndrome tentatively show similar performance for this group to typically developing participants on metonymy when constructed according to MA on the BPVS, whereas metaphor comprehension is severely affected. Our metaphor-metonymy cross-sectional study that traces performance from school-age to adulthood allows us to conclude that the timing of
developmental change is open to exogenous influences which are of great importance for future investigations.

Acknowledgements

We would like to thank all our participants, especially the children and their teachers for their continuing collaboration in our research. Special thanks to Anna Ferdenzi and Jo Van Herwegen for their assistance with recruitment and data collection. We are also grateful to Annette Karmiloff-Smith, Jill Hohenstein, and Michael Thomas for fruitful discussions. This research was supported by a KCL Seedcorn grant to G. R.

References


Received 2 March 2009; revised version received 7 May 2009
Appendix A

An example of a metaphor story: *flood* (meaning 'lots of people')

1. Stuart works at a museum. The museum is in the middle of town near a big river.

2. It is a small museum and not so many people come to the museum.

3. Stuart’s boss wants more people to come to the museum. So Stuart prepares a very special exhibition. Stuart’s boss tells lots and lots of people about Stuart’s exhibition.

4. It is Monday morning and Stuart is at home. Suddenly, the phone rings - it is Stuart’s boss. Stuart’s boss says. “You did it Stuart! There is a flood outside the museum.” Stuart runs to the museum to look. What does Stuart see?
An example of a metonymy story: *Robbie Williams* (meaning ‘CD’)

1. Kate and Anne are listening to music in Kate’s room. Kate has a lot of CDs with songs on.

2. Kate wants to play her favourite song to Anne. Kate looks for the CD with the song on.

3. But Kate cannot find the CD. She says, “Maybe my favourite CD is in another room”. Kate goes to look for the CD in the other rooms. Anne stays in Kate’s room.

4. After a while, Kate calls: “Come and look Anne! I found Robbie Williams in the lounge”. Anne goes to the lounge to look. What does Anne see?
Appendix B

An example of prompts for a metaphor story: flood (meaning 'lots of people')

PROMPT A1
Stuart’s boss says: ‘You did it Stuart! There is a flood outside the museum.’ Stuart runs to the museum to look. What does Stuart see?

PROMPT A2
Stuart’s boss says: ‘There is a flood outside the museum.’ Tell me about the museum and the flood!

PROMPT A3
Who is the flood?

An example of prompts for a metonymy story: Robbie Williams (meaning 'CD')

PROMPT A1
Kate calls: ‘Come and look Anne! I found Robbie Williams in the lounge’. Anne goes to the lounge to look. What does Anne see?

PROMPT A2
Kate calls: ‘I found Robbie Williams in the lounge.’ Tell me about Kate and Robbie Williams!

PROMPT A3
What is Robbie Williams?