Constructing Novel Iconic Signs Through Gesture

Dean Jones

A thesis submitted for the degree of Master of Arts at the University of Otago, Dunedin, New Zealand

February 2018
ABSTRACT

The origin of language has been a mystery for many years, with many possible theories offered as an explanation. One of the strongest theories states that human spoken language originated after the development of a gestural communication system. An important question in this theory is how these gestures developed, and how they became symbolic. Research within experimental semiotics has studied the process of simplification within emerging language systems, but most research has examined graphical systems and not gestures. The current research fills this gap by experimentally studying the process of simplification with iconic gestures. Ten pairs of participants were asked to use only bodily gesture to convey the meaning of a word to their partner. Concepts were repeated six times, allowing for the analysis of simplification over those representations. In addition, following Merola (2007) and Poggi (2008), four types of words (actions, animate creatures, natural objects, and artefacts) were provided in order to test the sorts of gestures that were used to convey iconic meaning. The results show that the number of component gestures used to convey the concept reduced over the first three iterations, after which they remained stable. In addition, there was a significant relationship between the type of meaning to be conveyed and the type of gesture employed. The thesis concludes examining the sorts of components that were employed and how simplification progressed in order to set up future research on the topic of emergent gesture-based communication systems.
## CONTENTS

Abstract........................................................................................................i

Table of Contents............................................................................................ii-iii

Acknowledgements............................................................................................iv-v

1 Constructing Novel Iconic Signs Through Gesture..................1-3
   1.1 Language Origin.................................................................4-11
   1.2 Gesture Research..............................................................11-22
   1.3 Experimental Semiotics......................................................23-27
   1.4 Aims and Hypotheses.........................................................28-34

2 Methodology...............................................................................................35
   2.1 Variables and Experimental Design.................................35-39
   2.2 Participants.................................................................39-42
   2.3 Experiment Set-Up and Equipment..................................42-43
   2.4 Procedure.................................................................................44-45
   2.5 Analysis..................................................................................45-47
   2.6 Example: *mosquito*..........................................................47-56

3 Results.......................................................................................................57-58
   3.1 Hypothesis 1.................................................................58-64
   3.2 Hypothesis 2.................................................................64-70
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3</td>
<td>Research Question</td>
<td>71-73</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Example: bag</td>
<td>74-93</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Example: mountain</td>
<td>94-107</td>
</tr>
<tr>
<td>4</td>
<td>Discussion</td>
<td>109-110</td>
</tr>
<tr>
<td>4.1</td>
<td>Hypothesis 1</td>
<td>110-117</td>
</tr>
<tr>
<td>4.2</td>
<td>Hypothesis 2</td>
<td>117-135</td>
</tr>
<tr>
<td>4.3</td>
<td>Research Question</td>
<td>135-138</td>
</tr>
<tr>
<td>4.4</td>
<td>Limitations and Future Research</td>
<td>139-142</td>
</tr>
<tr>
<td>5</td>
<td>Conclusion</td>
<td>143-145</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>147-156</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

The culmination of a thesis is always the work of not one person, but of all those who made contributions. Such is the case with this thesis, and I am grateful for the support of so many people.

In memory of Professor Jae Jung Song, I would like to dedicate this thesis to him. Before his passing last year, Jae consistently offered advice and encouragement on my thesis. Even before then he always strove to get the best out of me, as he did with many others. His belief in my academic aptitude helped me to work harder. At the times where my workload seemed insurmountable the memory of his words pushed me on.

I would like to acknowledge all twenty of the participants and thank them. Without their involvement this thesis would not have been possible to complete. I hope that the experiment was as fun for you as it was insightful for me.

I am particularly thankful for my supervisor, Dr. Hunter Hatfield. His knowledge of various fields of linguistics has proved invaluable throughout the duration of my research. I especially appreciate his rationality when I sent many emails of confusion and panic, as well as the engaging discussions on a variety of topics. Dr. Hatfield is not the only member of the department I would like to thank. Across my time as an undergraduate every member of staff has help guide and shape me, which continues even now. My sincerest thanks go to Dr. Anne Feryok, Dr. Moyra Sweetnam, and Dr. Simon Overall.
Another member of department that I would like to single out (and thank) is Dr. Karen McLean. She has endured my many questions, eclectic musings, and humoured me on many topics. I am very grateful for your support and wise advice!

I would like to extend my gratitude also to all members of the Otago Koto Ensemble (OKE) and O-Taiko. I might have let the stress get to me if I had not taken breaks from my thesis to play music with the lovely people in each of these groups. Their support and encouragement have helped fuel my determination.

My colleagues have also contributed to the completion of this thesis, no-one more so than my dear friend Elisha. Even while we are working on topics within different fields, Elisha was there to read my drafts when I needed another set of eyes and to be a sounding board for many of my ponderings. You already know you have my thanks!

I would also like to thank Ahmad who arrived not long after I began my thesis. Our infrequent discussions caused me to rethink my approach several times, ultimately, they caused me to engage in a more considered way of thinking. Your kindness and thoughtfulness are incredibly appreciated.

Finally, I would like to thank my closest friend and fiercest supporter – Kat Norton. Even though she spent months not understanding my ramblings about gestures and semiotics, her advice and encouragement have made it possible for me to keep going. There are many things I could mention to thank her for, but instead I will leave it at this – you have my heartfelt appreciation and gratitude.
1 Constructing Novel Iconic Signs Through Gesture

*I cannot doubt that language owes its origin to the imitation and modification, aided by signs and gestures, of various natural sounds, the voices of other animals, and man’s own instinctive cries.*

— Charles Darwin, 1871.

“The Descent of Man, and Selection in Relation to Sex”.

The theory of evolution states that humans are a product of progressive evolution that has continued for billions of years (Darwin, 1871). Somewhere along this evolutionary pathway, we developed the ability to use language; a communicative tool that is far more sophisticated and complex than discovered in any other species. In order to be able to claim that human language is exceptional in comparison to other communication systems, we must be able to define language and show where the division is between human language and animal communication. While we currently understand more about language than we ever have, there are still many mysteries to uncover. One of these mysteries is how language originated. Over the past few hundred years it has been a contentious topic, with many theories arising to explain how we came to use language (Berwick & Chomsky, 2016; Corballis, 2002; Deutscher, 2005; Dunbar, 1996; Falk, 2004; Heine & Kuteva, 2012; Knight & Power, 2012; Müller, 1996[1861]; Pinker, 1994; Rappaport, 1999; Tomasello, 1996; Ulbaek, 1998;). It is generally accepted that language emerged from sub-Saharan Africa around the middle of the Stone Age, assumedly contemporaneous with the speciation of Homo sapiens (Botha & Knight, 2009). A compelling account of
how language originated is the gestural theory of language origin. This account states that humans must have created language gesturally before developing speech to pair together with it (Arbib, Liebal, & Pika, 2008; Armstrong, 2008; Corballis, 2002, 2010, 2013; Premack, D., & Premack, A., 1983; Suddendorf & Corballis, 1997, 2007). One of the pieces of evidence for a gestural origin of language is the readiness of humans to construct signed language. There is consideration of the naturalistic studies on the formation of Nicaraguan Sign Language (ISN) (Kegl, 1994; Senghas, Kita, & Ozyurek, 2004; Senghas & Coppola, 2001). These theories, particularly the gestural theory, and the emergence of ISN are summarised in section 1.1.

Even in the modern day, gesture is a pivotal part of human communication; ranging from the paralinguistic gesture that accompanies speech to Sign Language gesture. It has been estimated that non-verbal communication comprises two-thirds of human communication (McDermott, 1980). Gesture is one of the main tools used in conveying meaning between people, and there has been a growing interest in researching linguistic gesture including Birdwhistell’s kinesics (1952) and McNeill’s work on gesture and thought (1992). Gestures are able to be classified into five different categories from gesticulation to sign languages. These categories have been captured by Kendon (1982) and McNeill (2013) in the Gesture Continuum. One of the points on the continuum is pantomimic gesture. Pantomime is a form of gesture that often occurs in isolation of speech, and is conducted by mimicking certain aspects of the signified object. The mimicked aspects are represented iconically through the gesturer’s motions. Iconicity is a term used in semiotic theory, which involves the triad of iconicity, indexicality, and symbolicity. While there is little research on iconicity in pantomime, iconicity in gesture has been investigated (Poggi, 2008).
summary of these types of gestural research, and an explanation of semiotic theory, are provided in section 1.2.

Experimental Semiotics (ES) experimentally tests aspects of language origin and the development of iconicity in language. Main areas of investigation of ES involve the progression of semiotics in language: how do iconic signs become symbolic? What process leads to a systematic use of language? Most of the current work (Galantucci, 2009; Galantucci and Garrod, 2011) in ES is focused on evidence of language creation using a pictographic mode of communication due to the ease of completely recording the novel communication system. The pictographic mode of communication is also referred to as a graphical modality.

Following these sections, the research aims and hypotheses are outlined in section 1.4. Due to gesture’s prevalence in communication this research will aim to provide ES research with a gestural experiment. A gestural experiment will expand ES research beyond solely graphically based experiments. As research of this kind has not yet been conducted, analytical tools will be formed based on Poggi’s work (2008) to specify iconic parameters of gestures.
1.1 Language Origin

The origin of language is a topic that has stimulated discussion for a long time, and still to this day we do not know how language truly originated. Even the famous Greek historian Herodotus had reported a story of the Pharoah Psammetichus conducting a language origin experiment (Histories, 2.2). The Pharoah took two newborn children and gave them to a farmer, instructing that no one was to speak a word within the hearing of the children. After 2 years, the children both started saying the word ‘bekos’ – the Phrygian word for bread. The Pharoah concluded that the Phrygians were an older civilisation, and perhaps the original language. The reliability of the Herodotus’ story might not be high, but the anecdote does show that questions surrounding language origin have existed for at least 3,000 years. Even now, the discussion surrounding the origin of language remains of interest, with many proposed theories and arguments (Atkinson, 2011; Berwick & Chomsky, 2016; Botha & Knight, 2009; Chomsky, 1996, 2004, 2005; Corballis, 2002; Deutscher, 2005; Dunbar, 1996; Falk, 2004; Fitch, 2004; Heine & Kuteva, 2007; Hopper, 1998; Knight, 1998, 2006, 2008, 2010; Knight & Power, 2012; Müller, 1996[1861]; Pika & Mitani, 2006; 2012; Pinker, 1994; Pinker & Bloom, 1990; Rappaport, 1999; Tomasello, 1996; Trivers, 1971; Ulbaek, 1998).

Before summarising the above plethora of theories of language origin, it is important to differentiate language and communication. This is important as all language is communicative, but not all communication is linguistic. While humans use language to communicate, there are other ways that we communicate that do not fit the current definition of language. This type of communication is commonly referred to as non-verbal communication, and can incorporate face gestures, body postures,
clothing vocal pitch, tone of voice, and gesticulation. Animal species also communicate through various means. Some communication is through vocal means, such as chirps, whistles, or growls, and some communication is through physical means, such as the bee dance, a canine baring its teeth, or a bird displaying plumage. We humans are the only ones that have a system of communication that can be classified in the way that language is.

Sapir defined language as “a purely human and noninstinctive method of communicating ideas, emotions, and desires by means of a system of voluntarily produced symbols” (1921, p. 7). However, the term language is not used to describe other forms of non-linguistic gestural communication between humans, nor is it used to describe the communication within other species. This is due to the definitions applied to language, stemming from design features itemized by Hockett (1960). Hockett stated thirteen design features that are shared across animal communication types, only three of which are particular to human language (Kirby & Smith, 2008). Kirby and Smith state that semanticity, productivity, and traditional transmission are the three features that separate human language from animal communication, but only when they appear in conjunction. Semanticity relates to the use of either arbitrary or non-arbitrary signals to convey a meaningful message. Productivity is “the capacity to say things that have never been said or heard before and yet to be understood by other speakers of the language” (Hockett, 1960, p.90). Traditional transmission is cultural transmission; it is transmitted extragenetically by means of teaching and learning. These features are based on Hockett’s 1960 work, and provide a framework with which to define language. Any of these three features can be found in communication when in isolation. However, only human language has all three features present at the same time. Kirby and Smith (2008) go on to state that these
three features are underpinned by the four subsidiary design features of arbitrariness, duality of patterning, recursion, and compositionality. These design features are separated from the distinctive features, as they can appear in non-human communications as well. All the features, both distinctive and subsidiary, are important parts of what is incorporated in complex human languages.

With the ability to define language by these design features, it is useful to see how they work in non-verbal communication and animal communication. All three features (semanticity, productivity, and traditional transmission) must be present within a system for this classification to work. We know that language already bears all three. Non-verbal communication does not usually operate linguistically by itself – it is normally paired with linguistic speech or linguistic gesture. This type of communication frequently accompanies speech, and individuals use hundreds of gestures each day¹ (Morris, Collett, Marsh, & O'Shaughnessy, 1979). This is noticeable in the use of gesture and facial movement during conversation but also in the use of ‘emoji’ in online, typed communication. The ‘emoji’ acts as supplementary information to inform the message recipient of the emotional context. For example the sentence “They have a cat” could be changed to “They have a cat 😊” (where the sender is happy about this), but it could also be changed to “They have a cat ☹” (where the sender is unhappy about this). Non-verbal communication does not

---

¹ It is commonly quoted that Mehrabian states only 7 percent of communication is verbal, with tone of voice accounting for 38 percent and body language account for the remaining 55 percent (Mehrabian & Ferris, 1967; Mehrabian & Weiner, 1967). These numbers are perpetuated in today’s academic world, but are inaccurately quoted (Lapakko, 2007). While it is hard to determine a precise percentage to show the prevalence of non-verbal communication, it is a safe assumption that non-verbal communication is widely used in conjunction with verbal communication to construct a social message.
operate linguistically as language does. Instead it is used to supply additional information during an interaction. Due to its co-dependence with language, it is not able to be defined as language when it is used in isolation.

Animal communication can be semantic, such as the alarm calls from various species of birds and monkeys (Cheney & Seyfarth, 1990; Evans, C., Evans, L., & Marler, 1993; Marler, 1955; Zuberbuhler, 2001). These alarm calls seem to be biologically transmitted, not culturally transmitted, and are not always productive. Some animals do use productivity in their communication, for example Gunnison's prairie dogs have been shown to incorporate information in their alarm calls relating to shirt colours of humans, as well as information relating to different sizes and shapes (Slobodchikoff, Kiriazis, Fischer, & Creef, 1991; Slobodchikoff & Placer, 2006). While the prairie dog calls are semantic and have productivity, they are not culturally transmitted. There is also evidence of some bird song being productive, such as that of the starling, willow warbler, and Bengalese finch (Eens, 1997; Gil & Slater, 2000; Okanoya, 2004), but these songs seem to only be semantic in a simplistic way; they only serve as either a way to attract mates or to repel rivals (Catchpole & Slater, 1995). Cultural transmission does seem possible in mammals, which can be seen through differing dialects in a species, such as within sperm whale colonies (Gero, Whitehead, & Rendell, 2016). Currently, non-human communicators have not been shown to incorporate traditional transmission, productivity, and semanticity. Therefore, these three features allow a distinction between language and both non-verbal communication and animal communication.

With the definition of language now more clear, it is even more of a wonder how language came to be. As it is such a unique and complex system, it is no surprise
that the point of origin has such a large number of theories surrounding it. The different theories can be separated into mutually exclusive subgroups based on shared assumptions (Ulbaek, 1998). The two main underlying assumptions are that: (a) language continually evolved, and did not just appear in the fully fledged version we have now (known as “continuity theories”)(Atkinson, 2011; Pika & Mitani, 2006; Pinker, 1994; Pinker & Bloom, 1990; Tomasello, 1996; Ulbaek, 1998), and (b) language is so unique that it must have appeared suddenly during evolution (known as “discontinuity theories”)( Berwick & Chomsky, 2016; Chomsky, 1996, 2004, 2005, 2010). A smaller third subgroup focuses on the origin and development of language as a social and cultural process (Knight & Power, 2012; Knight, 1998, 2006, 2008, 2010; Rappaport, 1999). Regardless of the subgroup that researchers align with, it is generally believed that language emerged from sub-Saharan Africa around the middle of the Stone Age, assumedly contemporaneous with the speciation of Homo sapiens (Botha & Knight, 2009).

Early speculative theories, named and published by Max Müller in 1861, claimed that language was originally: (a) an imitation of animal sounds (bow-wow theory), (b) natural emotional interjections caused by pleasure or pain (pooh-pooh theory), (c) based on objects’ natural vibrational frequency, similar to sound symbolism (ding-dong theory), or (d) caused by the desire to synchronize rhythmic labour jobs with alternating and opposing words such as ‘heave’ and ‘ho’ (yo-he-ho theory) (Müller, 1996[1861]). Another early speculative theory is the “ta-ta” theory, which claims that humans used the tongue to mimic manual gestures. This resulted in audible words (Paget, 1930). These theories are commonly seen as naïve and irrelevant among modern scholars (Firth, 1964; Stam, 1976). In recent years many theories have arisen, such as the mother tongues hypothesis (Fitch, 2004), the...
obligatory reciprocal altruism hypothesis (Trivers, 1971; Ulbaek, 1998), the gossip and grooming hypothesis (Dunbar, 1996), the ritual/speech co-evolution theory (Knight, 2010; Rappaport, 1999), the putting down the baby theory (Falk, 2004), and the grammaticalisation theory (Deutscher, 2005; Heine & Kuteva, 2007, 2012; Hopper, 1998).

The gestural theory of language origin is prominent, with a lot of theoretical and evidential backing (Arbib, Liebal, & Pika, 2008; Armstrong, 2008; Corballis, 2002, 2010, 2013; Premack, D., & Premack, A., 1983; Suddendorf & Corballis, 1997, 2007). This theory suggests that language, which is largely arbitrary, evolved from gestures that were non-arbitrary or iconic in nature (Armstrong, 2008). It relies on the neurological links between gesture and speech, for example the left cerebral hemisphere is specialised for motor selection for both speech and gesture (Kimura, 1993). This suggests that gesture most likely either developed in tandem with speech, or that it originated before linguistic speech. The gestural theory of language origin relies on the observation that nonhuman primates can also use gestures to communicate to some extent (Arbib, Liebal, & Pika, 2008; Premack, D., & Premack, A., 1983). Some of these gestures, such as the chest beating gesture of gorillas, are genetic. These types of genetic gestures are performed by primates who have never seen another primate perform that gesture. It has been suggested that, due to the increased use of tools, it was no longer convenient to use hands for gesture and speech had to develop (Corballis, 2002). Corballis (2010) proposes several pieces of evidence to support the gestural origins of language. They are:

1) the use of the hands as the more natural way to depict events in space and time; 2) the ability of nonhuman primates to use manual action flexibly and
intentionally; (3) the nature of the primate mirror system and its homology with the language circuits in the human brain; (4) the relative success in teaching apes to communicate gesturally rather than vocally; (5) the ready invention of sophisticated signed languages by the deaf; (6) the critical role of pointing in the way young children learn language; and (7) the correlation between handedness and cerebral asymmetry for language. (p. 1)

While some researchers believe that language developed in one sudden mutation (Chomsky, 2010; Tattersall, 2012), Darwin’s theory of evolution presumes that all evolutionary development is gradual rather than sudden (Darwin, 1859; Pinker & Bloom, 1990). Using a more visible form of communication during language development allows for people to establish meaning before moving to speech. This is especially so due to the ease of using iconic forms in gesture in comparison to abstract forms of representation in speech. Abstract forms of language are symbolic, gradually developing from iconic forms (Galantucci & Garrod, 2011). The gradual development of symbolicity supports the idea of languages gradually developing, and does not support the theory of a sudden mutation. Following the development of our original gestural language, Corballis hypothesises that speech developed with paralinguistic gesture (Corballis, 2002). This is supported by the current use of paralinguistic gesture alongside speech, as well as the shared neurological facilities for both speech and gesture (Kimura, 1993).

Instances of the development of a signed language have been researched in recent years (Senghas & Kegl, 1994; Kegl, 1994; Kegl, Senghas, & Coppola, 1998; Pyers & Senghas, 2006; 2005; Senghas & Coppola, 2001; Senghas, Kita, & Ozyurek, 2004). An example of signed languages developing is Nicaraguan Sign Language. In
the nineteen seventies, there was a new school founded in Nicaragua especially for deaf people. Before the school was established, there had been no conventionalised language form for deaf people, though they would have used home-signs. These home-signs varied in complexity and form for each person (Senghas et al., 2004). The creation of the school drew deaf people together for the first time in large numbers, and they began creating a language together. This sign language was titled “Lenguaje de Signos Nicaragüense” (LSN), but after the development of the language by children, it was later renamed “Idioma de Señas de Nicaragua” (ISN). The chance to document a naturally developing language from its early stages is rare, and the development of ISN has shed some light on the process of gestural language evolution. It demonstrates that humans will readily adopt a gestural form of language.

To speculate further on a gestural origin, modern research on gesture is discussed in section 1.2. In particular the role of iconicity in gesture is raised.

1.2 Gesture Research

As mentioned, gesture can be used in both a linguistic way, such as in sign languages, and non-verbally. Gestural research is still a young scientific field, with researchers examining various aspects of gesture such as cultural gesture (Ting-Toomey, 1999), pragmatic gesture (Kendon, 1995; Müller, 2004), and sign language gesture (Goldin-Meadow & Brentari, 2017). While there is a lot of focus in research on paralinguistic gesture (how speech and gesture work together to convey meaning), there is also research on signed languages (Kegl, 1994; Kegl, et al., 1998;
Constructing Novel Iconic Signs Through Gesture

D. A. Jones

Klima & Bellugi, 1987; Polich, 2005; Pyers & Senghas, 2006; Senghas & Coppola, 2001; Senghas et al., 2004). This section details gesture and considers some of the research associated with non-linguistic gesture, such as paralinguistic gesture (Birdwhistell, 1952; Kendon, 1982; McNeill, 1992, 2013) and pantomimic gesture (Donald, 1991; 2001; Tomasello, 2008; Zlatev, 2008, 2014; Żywiczyński, Wacewicz, & Sibierska, 2016).

One of the first major linguistic investigations into the parallels and interplay between speech and gesture was conducted by Ray Birdwhistell. Throughout his essays on the matter, he developed the term “kinesics” and compared all minor body movements (kines) to minor vocal tract movements (phones) (Birdwhistell, 1952). One large issue with this consideration of paralinguistic gesture is that the kines must be precisely and laboriously transcribed, as there are often multiple, co-occurring sources of movement in the body. His work has been considered as the general study of non-verbal communication by some researchers (Duncan, 1969; Knapp, 1972; von Raffler-Engel, 1980), while others consider it as an approach to the study of verbal behaviour (Argyle, 1975; Scherer & Ekman, 1982; Scherer & Giles, 1979). Birdwhistell’s work has been a good beginning for research on gesture. However the process of analysis required for his theory of kinesics is extensive. Laboriously documenting all forms of gestural movement is useful for a detailed account of gesture use. Regardless of its usefulness, a researcher would have to dedicate a sizeable amount of time to analyse even 10 seconds of recorded gestural data.

Another issue with this type of analysis is that the components that contain meaning can be obscured. This is less true of paralinguistic gesture and is more applicable to gesture that is used by itself, such as sign language gesture or pantomimic gesture.

When a person makes a gesture, for example a ‘thumbs up’ gesture, Birdwhistell’s
analysis would require notation of all motions corresponding to the component of meaning (the thumb). This could include motions of the arm, torso, or any other body part that was incidentally involved. With the invention of motion sensor technology, Birdwhistell’s analytical ideas can be achieved in an easier and quicker fashion.

David McNeill primarily focuses on the interaction of gesture and speech in the representation of thought. As with Birdwhistell, McNeill’s work is based in paralinguistic gesture. He has detailed previous work by Kendon (1982) and named it the “Gesture Continuum” (McNeill, 2013). This continuum demonstrates the different types of gesture, as previous work had been using the term ‘gesture’ holistically, detrimentally ignoring distinctions between paralinguistic gesture, non-linguistic gesture, and linguistic gesture (such as sign languages). The continuum is:

Gesticulation -> Language-like Gestures -> Pantomimes -> Emblems -> Sign Languages

As McNeill states, gesticulation is the specific area of gestures that he is primarily interested in, and he uses the word gesture “specifically to refer to the leftmost, ‘gesticulation’ end of the spectrum” (McNeill, 1992, p.37). Gesticulation refers to any paralinguistic gesture and rarely occurs without a spoken accompaniment. Language-like gestures are used during sentence fragments to express that which is left unspoken. McNeill uses the example “the parents were all right, but the kids were [gesture]”, where the gesture replaces an adjective (McNeill, 1992, p.37). Pantomime is an interesting type of gesture, as it refers to gestural action that depict objects or actions usually in an iconic way. These actions might be accompanied by non-linguistic, onomatopoeic noises, for example when pantomiming an aeroplane, a person might mimic the sound of a plane at the same time. Pantomime is different to gesticulation in several ways, but an important
distinction made by McNeill is that pantomimes can create sequences of gestures while gesticulation cannot. Emblems are well-formed gestures that carry meaning, such as the Western hand sign for ‘OK’ (made with the index finger and thumb joining). If any other finger was to touch the thumb instead of the index finger the meaning would be changed. McNeill states that the position of emblem and pantomime within the continuum is arbitrary. Finally, Sign Languages are wholly linguistic gestural systems with segmentation, compositionality, a lexicon, a syntax, distinctiveness, arbitrariness, standards of well-formedness, and a community of users (Klima & Bellugi, 1987; McNeill, 1992). McNeill’s work focuses on gesticulation, which he simply refers to as ‘gesture’, and his analytical framework is designed primarily for gesticulation, rather than other types of gesture used within language-like gestures, pantomime, emblems, and sign languages. In this thesis the two terms ‘gesticulation’ and ‘gesture’ are not used synonymously: ‘gesture’ relates to any type of meaning-bearing movement.

When considering the gestural origin of language, the Gesture Continuum provides an interesting thought experiment. Gesticulation and language-like gestures on the left hand side of the continuum are paralinguistic. Gesticulation is used simultaneously with speech, while language-like gestures act as gap-fillers in speech. Both emblems and sign languages must be developed before use. This is due to their use of more abstract and symbolic forms, which can make communication difficult if the other person does not understand what signified object you are gesturing.

Pantomime is the only type of gesture on the continuum that potentially could have been used to develop meaning during the genesis of human language. The reason pantomime would have been used to establish meaning is that it is derived from
bodily mimesis (Donald, 1991; Zlatev, 2014). Żywiczyński, Wacewicz, and Sibierska (2016) define pantomime as:

A communication mode that is mimetic; volitional and representational; non-conventional and motivated; multimodal (primarily visual); improvised; using the whole body and the surrounding space rather than exclusively manual and stationary; holistic and non-segmental; communicatively complex and self-sufficient; semantically complex; displaced, open-ended and universal. (p. 1)

In contrast, non-pantomimic gesture is frequently regarded as involving movements of the hands. Pantomimic gesture, however, can incorporate the whole body to mimic what it is representing. In everyday life pantomime is not used by individuals as commonly as gesticulation and has not been a focus of gesture research (see Arbib, 2012; Donald, 1991; Tomasello, 2008; Zlatev, 2008). Żywiczyński et al. raise the valid issue that the field of pantomime research suffers from a lack of theoretical explanation and exploration for empirical purposes. Without a solid theoretical explanation and exploration, pantomime research is “more intuitive than systematic” (Żywiczyński et al., 2016, p. 1). It is also worth noting that ‘pantomime’ refers to whole events and can be difficult to segment into clear parts, especially as there are “no self-apparent onsets and terminations in the stream of movement” (Żywiczyński et al., 2016, p. 8).

To mimic something is to recreate an aspect of the object. This is more succinctly described using semiotic terminology, such as iconicity and symbolicity. The semiotic theory of the sign is briefly summarised here. For communication to be possible, interlocutors must produce signs. These signs are capable of being produced on any medium involving the senses, but signs are commonly vocal,
gestural, or graphical. Signs are not only used for interpersonal communication – it is through signs we interpret and understand the world. A sign can be iconic, indexical, symbolic, or any combination of these three (Nöth, 1995). If a sign is iconic, then it resembles an aspect of the signified object. For example, a person might ‘flap’ their arms to represent a bird. While the action is not a bird, it is bird-like and can be easily understood. Iconicity is useful in communication, as interlocutors do not need to have a pre-established, or grounded, understanding of the sign for it to make sense.

Iconicity is also present in onomatopoeic speech, for example 'bang' is iconic of the noise it represents. If a sign is indexical, it points to or indicates the intended object. For example, smoke is an index of fire; not only do we know that fire causes smoke, they are also spatially connected (i.e., in general smoke does not appear where there is not fire). Because of this spatial connection, we know that the smoke will lead us to its source. Another example of an index is pointing at an object, which draws attention to the object, allowing all interlocutors to share the same understanding of what is signified. Spoken language also uses indices, such as the words here and there.

If a sign is symbolic, the signifier generally bears no resemblance to the signified and the meaning is understood through cultural convention. For example, many (if not all) orthographic systems are symbolic - English written letters are not iconic of anything else, they are not indexical, and they must be learned to understand the convention. The majority of words in all languages are symbolic: cat abstractly represents a cat as much as the Japanese word for cat, neko does.

---

2 There is discourse around scripts, such as Chinese, which are believed to have been iconic or pictographic originally. However, in the present day these scripts have abstracted from earlier forms and are culturally embedded symbols.
In this thesis, the term ‘iconicity’ is used to state the resemblance a signifier has to the signified object or action. This can be a resemblance of shape, size, motion, or corresponding action performed on the object. While the term ‘iconicity’ is used, it is important to note some differences in the use of the word ‘iconicity’ in relation to gestural research. ‘Indexicality’ is also a suitable word to describe the relationship between signifier and signified. As Streeck (2008) details, iconicity states similarity between both the signified and signifier. This is to say, both the signifier (X) and signified (Y) must bear a likeness to one another (X=Y, Y=X). However, the signifier is only a representative for the signified, and only gives indexical clues resembling the signified (X=Y, Y≠X). Using two hands, interlocked at the thumbs, to represent a bird or a butterfly is a one-way resemblance; a bird or butterfly does not resemble the hand gesture, but the hand gestures do resemble aspects of a bird or a butterfly.

Pantomime is a type of gesture that employs mimicry of signified concepts. In other words, pantomime uses iconic forms to help communicators understand each other. While there is a lack of research specifically focussed on pantomime, research on iconicity within gesture is not novel, with several researchers conducting experiments (Kendon, 1980; 1982; Magno Caldognetto & Poggi, 1995; Mandel, 1977; Merola, 2007; Poggi, 2007), as well as discussion surrounding what iconicity means when pertaining to gesture (Ekman & Friesen, 1969; Mandel, 1977; McNeill, 1992; Müller, 1998; Streeck, 2008). Poggi has researched gesture since the 1980s, and has developed theories relating to gesture, but especially relating to iconic forms within gesture. Poggi (2008) defines gestures as “any movement performed by hands, arms, or shoulder” (Poggi, 2008, p. 46). While pantomimic gestures are capable of being performed by the whole body, communicative gestures are commonly performed by the hands, arms, or shoulders.
Poggi claims that all gestures can be classified within six different parameters: semantic content (“information on the world, on the Sender's mind, on the Sender's identity”), goal source (“individual, biological, or social”), level of awareness (“conscious, unconscious, or tacit”), relationship to other signals (“autonomous vs. co-verbal”), cognitive construction (“codified vs. creative”), and gesture-meaning relationship (“motivated (natural or iconic) vs. arbitrary”) (Poggi, 2008, pp. 47-49). These parameters are well suited to describe gestures that are used paralinguistically. Two of the six are also applicable to gesture not conducted in tandem with speech, namely cognitive construction and gesture-meaning relationship. Cognitive construction is separated into either codified or creative gesture. A codified gesture is one that has been pre-established within the mind of the gesturer, and it operates as part of a stable gesture-meaning pair. This is not limited to culturally specific gesture, but instead incorporates all gesture-meaning pairs that a communicator has developed throughout their lifetime. In other words, a codified gesture is a symbolic gesture.

A creative gesture occurs when a gesturer lacks a pre-existing gesture with which to communicate an idea. This type of gesture is constructed on an implicitly understood set of rules. For example, an object can be referred to by various means: pointing at it, imitating its shape or movements, or imitating the actions performed on it. To simplify, creative gestures can either be deictic (object indicated through pointing) or iconic. Poggi describes a creative iconic gesture as “a creative gesture that reproduces some perceivable aspects held by or linked to the meaning it conveys” (Poggi, 2008, p. 49). A gesture-meaning relationship can either be motivated or arbitrary. If the relationship is motivated, there is some link between signifier and signified that is apparent to an outside observer that has never seen the
motivated sign before. If the relationship is arbitrary, there is no apparent link
between the signifier and signified. Poggi separates motivated gestures into iconic or
natural gestures. As mentioned, an iconic gesture reproduces some perceivable
aspects of the signified object. A motivated gesture that is natural derives its meaning
from mechanical determinism; actions produced by physiological impulses are
natural. For example, the shaking of a raised fist to express happiness or celebration
is not iconic. Instead it is linked to the physiological activation produced by emotions
(Anolli, Agliati, Chinnici, & Crippa, 2002). An important link between these two
parameters is that an arbitrary gesture must be codified to be understood, while a
creative gesture must be motivated (Poggi, 2008).

As mentioned previously, within the creative gesture parameter there are
various types of iconicity. A clearer explanation for each type is useful when drawing
distinctions. These types of iconicity have been proposed as a model for the
generation of creative iconic gestures (Magno Caldognetto & Poggi, 1995; Poggi,

1. the REFERENT’S SHAPE: to mean “mountain” you can outline a conic shape,
for “brioche” your hand forms a rounded shape
2. the REFERENT’S LOCATION: to mean “sign”, you draw a rectangular shape up
over you (sic) head; for “mountain” you can point far away
3. the REFERENT’S TYPICAL ACTIONS: to mean “bird” you move hands like wings;
for “lightning” you depict its zigzag trajectory
4. an AGENT’S ACTION with the referent: to convey “mountain” you pretend
climbing or skiing

(Poggi, 2008, p. 52)
The type of iconicity borne in the above four types is representational. The representation is only capable of conveying aspects of the signified. Physical entities are easier to represent iconically. Abstract concepts, such as love, science, or green, are more difficult to represent iconically. Some natural phenomena are also difficult to represent iconically through gesture. For example, wind is a natural force that interacts with objects, but has no physical mass that can be represented by itself. While these types of words are difficult to represent visually, they are not impossible. To convey the meaning of these words, participants resort to a Medium Referent (MR), which is something that can be represented through hand movements but makes inference to the mimed referent (or ‘Target Referent’, TR). This inference allows a bridge to form between a physical representation and abstract meaning. MR and TR are defined in four primary ways in Poggi (2008), as shown below:

1. **A Cause-Effect relationship**, where MR, with respect to TR, can be
   a. a cause of TR: to mean “surprise” subjects represented offering a gift – a possible cause of surprise;
   b. an effect of TR: to mean “wind (sic), subjects mimed waving hair or leaves; to mean “noise”, they displayed an annoyed face;
   c. an agent who is a source of the TR: for “science”, subjects mimed the actions or the look of a scientist;
   d. a location seen as a source of the TR: for “idea”, subjects pointed their index finger to their head.
2. **An Object-Function relationship**: to represent “noun”, a subject depicted a label on her breast.
3. **A Class-Prototype relationship**: to mean “dictatorship”, a subject mimed Mussolini.
4. A relationship of **OPPOSITION** and its **NEGATION**: to mean “democracy”, a subject mimed the typical posture of Mussolini, seen as a prototype of “dictatorship”, the opposite of “democracy”, and then denied this. To mean “freedom”, a subject pretended to be behind the bars of a prison (the opposite of “freedom”) and then to come out of it.

(Poggi, 2008, p. 53)

The framing of iconicity in this way allows for documentation of gestures, without the researcher having to micro-analyse the smallest components involved in gesture. Using these parameters researchers have investigated iconicity within gesture. These four parameters were observed in a study by Merola (2007) on iconic gestures used by school teachers while instructing new vocabulary. Table 1 shows the results from the study.

**Table 1**

*Results from Merola (2007)*

<table>
<thead>
<tr>
<th>Word Type</th>
<th>Shape</th>
<th>Location</th>
<th>Referent Action</th>
<th>Agent Action</th>
<th>TOT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Action</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.</td>
<td>%</td>
<td>n.</td>
<td>%</td>
<td>n.</td>
</tr>
<tr>
<td></td>
<td><strong>18</strong></td>
<td><strong>100</strong></td>
<td><strong>18</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Artefact</strong></td>
<td>11</td>
<td>23</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td><strong>Natural Object</strong></td>
<td>10</td>
<td>83</td>
<td>2</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td><strong>Animate</strong></td>
<td>2</td>
<td>18</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td><strong>TOT</strong></td>
<td>23</td>
<td>18</td>
<td>3</td>
<td>17</td>
<td>46</td>
</tr>
</tbody>
</table>

*The highlighted cells indicate the highest percentage of an iconic parameter's use for each concept type.*

Each of the four concept types showed a preference for particular parameters. Actions were always demonstrated through iconic representations of an agent’s action (100%). Artefacts were also demonstrated through agent’s actions (56%), natural objects were most commonly demonstrated by iconic depictions of shape (83%) and animates were most commonly demonstrated by depictions of a referent’s action. For example, natural objects (such as *sun* or *tree*) were more likely
to be described through iconic shape, while animates (such as dog or mouse) were more likely to be portrayed through referent action. Poggi concludes that “[t]hese results confirm that the ontological type of the meaning to be conveyed also affects the choice of the features to be represented by the hands” (Poggi, 2008: p. 55). Poggi continues:

> [e]very time one has to create an iconic gesture to convey a concrete referent one represents some distinctive aspects of it: the Referent’s Location, Shape and Actions, and the Actions performed on/through/about it. The aspects selected to be represented manually highly depend on their degree of distinctiveness, on their ease of depiction, and on the ontological type of the target meaning – whether action or entity, animate or inanimate, natural object or artefact (p. 55).

In this section, several topics of gesture were discussed such as Birdwhistell’s kinesics, McNeill’s paralinguistic gesture and the gesture continuum, pantomime, semiotic theory, and Poggi’s iconicity within gesture. As iconicity precedes symbolicity within the theory of the sign, it is reasonable to assume that the original form of human language utilized iconicity. Pantomime is a form of gesture that employs mimicry to convey meaning. This is achieved by the mimed gesture bearing an aspect of iconicity, or resemblance, to the signified concept. Due to the use of iconic forms in pantomime, it would make sense to conduct experimental research using pantomime to investigate the origins of language. To create such an experiment, research from experimental semiotics is discussed in section 1.3.
1.3 Experimental Semiotics

An alternative approach to studying language origin is to document the birth of new language systems in an experimental setting. While the creation of novel languages has been documented in naturalistic studies (Kegl, 1994; Kegl et al., 1998; Senghas et al., 2004; Senghas & Coppola, 2001), experimental studies allow for recreatability and specific exploration of key aspects of what is involved in language creation. These studies fall under the category of Experimental Semiotics (ES), a term coined by Galantucci (2005). The field is similar to Experimental Pragmatics (EP), which focuses on understanding how the many actions involved in spoken language work and interact. ES looks at human communication in general, not just spoken human communication. The field particularly looks at the role of semiotic elements within communication, such as iconicity, indexicality, and symbolicity. ES adopts the same core assumptions as EP, such as

the assumption that communication is a realtime social process which must be understood at the level of dyadic interactions (Pickering & Garrod, 2004) ...

that communicative interactions are embodied in the physical world (Goodwin, 2000) and embedded in fairly rich socio-cognitive contexts (Brennan & Clark, 1996; Hutchins, 1995; Krauss & Glucksberg, 1977; Suchman, 1987) (Galantucci, 2009, p. 394).

Even with the same core assumptions, ES varies in two primary ways. The first of these has already been mentioned; ES focuses on human communication in general. This means that physical and graphical communication are investigated in addition to verbal communication. The other difference is stated clearly in Galantucci (2009): "Experimental semiotics studies the emergence of new forms of
communication; experimental pragmatics studies the spontaneous use of preexisting forms of communication such as spoken English” (pp. 394-395).

Galantucci and Garrod (2011) summarises the main varieties of studies in ES as semiotic referential games (Healey, King, & Swoboda, 2004; Healey, McCabe, & Katagiri, 2000; Healey, Swoboda, Umata & Katagiri, 2002; Healey, Swoboda, Umata, & King, 2007; Garrod, Fay, Lee, Oberlander, & MacLeaod, 2007; Garrod, Fay, Rogers, Walker, & Swoboda, 2010), coordination semiotic games (de Ruiter, Noordzij, Newman-Norlund, Hagoort, & Toni, 2007; Galantucci, 2005, 2009; Noordzij, Newman-Norlund, de Ruiter, Hagoort, Levinson, & Toni, 2009; Scott-Phillips, Kirby, & Ritchie, 2009;), and semiotic matching games (Kirby, Cornish, & Smith, 2008; Roberts, 2008, 2010; Selten & Warglien, 2007). In referential games, participants typically have a pre-established set of referents. In coordination games, participants are free to discover any referent that helps them succeed in the given task. Both the referential and coordination games are useful for studying how communication systems emerge, but do not allow much study of more complex structures emerging in the communication system. This is because participants are often able to complete the tasks using simple systems. Semiotic matching games typically provide participants with a limited set of communication forms and referents. Experiments using semiotic matching games can show the emergence of compositionality, as well as the role of social dynamics.

Early research in ES was conducted by Healey, McCabe, & Katagiri (2000). Participants created a novel communication system with a graphical medium, but were not allowed to use letters or numbers in their interactions. They were asked to draw a representation of music that they had listened to, or to draw an abstract
concept for a partner to recognise. The field of ES has produced some innovative ways of conducting similar experiments using a graphical interface. For example, Galantucci (2005) implemented an electronic drawing pad in which the cursor consistently moved diagonally. This made direct graphical representation difficult for participants, and they were forced to be more creative during their communications. These studies found that over the course of the interaction, the graphemes used became more abstract and simplified. The importance of negotiation in developing a new communication system has also been highlighted. Participants were more successful in completing the tasks if they had engaged in graphical feedback. Participants that were assigned the task of overseers were poorer at understanding the graphemes compared to participants that were directly involved in the communicative task.

The previous research in ES has revealed several factors involved in creating a new communication system. Some of these have already been mentioned, such as negotiation which is a key component in developing a system, and signs which develop from iconic to abstract after their meaning is clear to both interlocutors. It has also been found that even early in a language’s creation there can be a large degree of combinatoriality (Galantucci, Kroos, & Rhodes, 2010). This was especially the case when the grapheme faded rapidly, compared to a slowly fading grapheme. The grapheme was designed to have different rates of fading to replicate how quickly the speech signal fades. Roberts (2008, 2010) observed differences in a newly-acquired artificial language’s use depending on two social factors: how frequently participants interacted with each other, and whether they were competing. The more frequently the participants interacted with each other, the more they identified subtle linguistic cues made by their partner. When frequency of interaction was paired with
competitiveness, linguistic divergence occurred, which shows that “when human interactions are both conflictive and frequent, linguistic divergence can occur at a very fast pace” (Galantucci & Garrod, 2011, p. 8). These discoveries are primarily for dyadic interaction, but also pertain to interaction in communities.

Research also shows there is a difference between dyadic pairs and larger groups when constructing a communication system. For example, Fay, Garrod, and Roberts (2008) showed that symbols created by a group, or community, were able to be matched to their referents at a faster and more accurate rate than symbols produced by pairs.

A great benefit of ES studies that use a graphical interface is that it allows the experimenter to have control over many features of language use. By removing the spoken and physical elements, a more focused look at the development is possible. Another benefit is that all interactions within the novel communication system are fully recorded, allowing documentation of the whole system. However, while there are benefits to observing communication systems developing over a graphical interface, it would be advantageous to observe novel communication systems developing through physical gesture. All of the previously mentioned ES research has been based in graphical communication only. A gesturally-based experiment would offer a greater wealth of data when compared with a graphically based experiment due to: (a) the close relation between speech and gesture (Kimura, 1993; Newman, Bavelier, Corina, Jezzard, & Neville, 2001); (b) the highly frequent use of gesture in communication; and (c) the greater variety of negotiation that can occur between participants who are face-to-face. As mentioned, research has found that the communicative activities were more successful when participants actively negotiated
the meaning of signs. It was noted in Galantucci (2005) that, when participants could view each other's virtual character (avatar), negotiation occurred. Participants moved their avatars up and down for yes (similar to nodding) and left and right for no (similar to a person shaking their head). This type of negotiation is important for verifying that both participants understand the communicative gestures and to minimize misunderstanding. In developing a communication system, “the increased simplification and evolution of graphical forms depended crucially on communicators' ability to give graphical feedback to their partner” (Galantucci & Garrod, 2011, p. 2). In face-to-face communication, it is rare to use only yes or no. Instead there are many words and movements used to indicate intermediary terms (maybe, sort of, etc.). The variety of negotiation techniques that exist in face-to-face communication would allow for more understanding between the participants. A novel communication system would benefit from more freedom to identify precise meanings of communicative gestures and establish mutual understanding. In addition, a gestural experiment would help to give evidence to the theory of a gestural language origin.

This section has discussed the field of experimental semiotics as well as the current research paradigms that are used. These paradigms are centered around researching semiotic development in a graphical modality, i.e. through pictographic forms of language. An alternative approach was also suggested; conducting ES research using a gesturally-based experiment instead of a graphically-based one. Section 1.4 expands on this and presents the aims and hypotheses of the current research.
1.4  -  **Aims and Hypotheses**

Having considered various research that incorporates gesture and as well as the graphically based research in ES, it will now be argued that research similar to previous work in ES could be conducted, using gestural (or manual) communication instead of graphical communication. This gestural approach to ES research is motivated by two main reasons.

The first motivation for the current research is that experimentally recreating a more natural context would show an even greater degree of authenticity within the language creation process, as well as providing the same recreatability that ES provides. A gestural experiment might also reveal more about the gestural origin of language than to a graphical experiment would, especially as it is argued that spoken language first originated from gesture and there is no evidence suggesting that language began graphically. Investigating the invention of a novel gestural communicative system is similar to previous naturalistic research on the development of Nicaraguan Sign Language (ISN) (Kegl, 1994; Senghas et al., 2004; Senghas & Coppola, 2001). Through the development of ISN, gestural language was seen to develop from the ground up in an entirely natural context. Due to the natural context, this situation provided little evidence for “the extent to which key features of natural languages ... arise from general principles of human communication rather than specific characteristics of those languages” (Galantucci & Garrod, 2011, p. 4). To investigate the development of iconicity in human communication it is important to be able to locate general patterns instead of considering only one language’s development specifically.
The second motivation for the current research is that it is unlike ES research which is based on graphical modality, limiting negotiation between participants. Though ES research is useful and has clarified processes involved in the development from iconicity to symbolicity, the research field is still young and further research is required to solidify current findings, as well as to expand their applicability.

As negotiation is crucial in the development of a novel communication system, experiments based on face-to-face communication would allow participants a freer range of negotiation techniques. Communicating in person could entail either spoken and/or gestural means of communication.

To observe the development and progression of iconicity within gestures, a solely gesture-based experiment is proposed. Speakers of a language will often use paralinguistic gesture but will less often gesture as a means of communication by itself. Removing the use of speech becomes a creative limitation, as in Galantucci (2005), which should cause participants to think carefully about how they wish to communicate gesturally.

An experiment that encourages participants to engage in pantomimic forms would be advantageous, as pantomime naturally engages in mimesis. This means that pantomime uses iconic forms frequently, as mimetic gesture represents some salient aspect of the signified object.

The experiment is designed to draw the gestures away from a participants’ cultural prior knowledge of gesture. This is to allow for participants to construct a novel language without the scaffolding of their previously-existing spoken knowledge base, which would cause difficulty in determining to what extent their ‘novel language’ is truly novel.
To summarize, experimental semiotics has largely utilized experiments based in the graphical modality. It would be advantageous to conduct experiments which are similar in nature to ES, but use a different modality. Gestures are commonly used in communicative practice, and the gestural (or manual) modality will produce useful comparative data.

Previous studies in gestural language development focussed on ISL in naturally occurring contexts, while this experiment will allow for observations to be made in an experimental and recreatable environment. This thesis aims to be a starting point for future researchers wishing to compare the development from iconicity to symbolicity in graphical experiments and gestural experiments.

In order to relate the previously-discussed graphical based experiments to gesture-based experiments, similarities between the two communicative styles must be drawn. Using the notion of kinemes, or other micro-units of gesture, in comparison to graphemes is insufficient, as the small units relate to their specific medium. The most important commonality between modalities, whether gestural, graphical, or spoken, is that they all refer to something in the world. This relationship between signifier and signified has core properties which are similar between each modality, even though the end product varies. Graphical representations originate with some degree of iconicity, which is progressively simplified and abstracted (Galantucci, 2005; Healey, McCabe, & Katagiri, 2000; Healey, Swoboda, Umata, & Katagiri, 2002; Healey, King, & Swoboda, 2004). Similar progressive development has been observed in manual gestures (Kegl, 1994; Senghas & Coppola, 2001; Senghas et al., 2004). As has been mentioned, pantomimic gesture uses mimetic forms to represent an aspect,
or multiple aspects, of a signified object iconically. This makes it the ideal form of gesture to use for investigating the development of iconic forms.

With the previous work on iconicity within gestures (Merola, 2007; Poggi, 2008), a framework for analysis of pantomimic gesture may easily be created (detailed in Chapter 2).

This study will investigate two research hypotheses, and one central research question. The hypotheses are as follows:

(1) that participants will simplify their iconic gestural components following the establishment of mutual understanding;

(2) participants will primarily choose to represent action word types by an agent’s action, artefacts by an agent’s action, natural objects by shape, and animates by a referent’s typical action.

The central research question is:

*How do iconic manual gestures become simplified?*

To test the research hypotheses and the research question, the experiment for this study requires participants to perform a task similar to that in the game of charades. While the experimental design is listed in full in Chapter 2, the general outline of procedures is described here. Ten pairs of participants were asked to convey 120 concepts to each other (60 concepts each) through gesture alone. These were made up of 20 individual concepts, with 6 iterations for each concept. Participants took it in turns to be ‘gesturers’ and ‘guessers’. After gesturing 10 concepts, participants swapped roles. After the experiment was complete, a short
interview was conducted with each pair to discover their rationale for some of the gestures they had chosen to perform.

As the field of ES focuses more on general novel human communication, instead of spontaneous use of pre-existing forms, it is important to consider to what extent adult users of a language can create entirely novel forms, separate from pre-existing forms. For this reason, verbal forms of language will be discouraged. Physical communication is closely tied with verbal language and often accompanies it. Because of this relationship between gesture and speech, the creation of an ES experiment that focuses on physical communication alone requires creative limitation (Galantucci, 2005). This limitation should draw the physicality away from a participants’ prior knowledge of cultural gesturology. It would be impossible to entirely limit a participant’s cultural knowledge. However reducing the mode of communication from primarily spoken to solely gestural will prevent participants from easily using pre-existing cultural forms. As participants will all be capable of spoken language, it is assumed that their use of solely gestural forms of communication will be infrequent.

As the research is situated within ES, an examination of gesture on a micro level is not useful. Transcribing every movement from all relevant body parts for a gesture would not only be excessively time consuming, but also an analysis that includes so many micro-movements may end up obscuring the iconic form that is intended. This is due to iconicity being observable within larger chunks, or components, of gesture only. For example, if the participant tried to iconically represent a concept such as dog they might choose to mimic the typical actions and shape of a dog. They could choose to mimic a dog by getting on to their knees and
then miming a barking motion. In order to fully document every movement involved a researcher would need to note the movements of the legs, arms, torso, head, etc. Iconicity would be obscured at such a microscopic level, rendering such analysis useless for this type of research. Poggi (2008) analyses gestures by the movements of the arms and hands only, limiting how much there is to analyse. This is possible with gesticulation and language-like gestures as they are paired with speech, and frequently use only the arms and hands. However, pantomime is a form of gesture that can use the whole body in order to mimic the referent. Therefore analysing only the motions of the arms and hands would be insufficient to capture the full body gestures that pantomime requires. In order to capture the iconicity that is present in the pantomimic gestures the notion of a gestural component will be used.

A component in this research is similar to a syntactic construction. It is a compact bundle of gestures that relates to one semantic meaning. A component is subjectively decided by the researcher when it is apparent that a participant has gestured a new meaning, for example pinching fingers to each other to depict ‘small’ and ‘crawling’ a finger along the forearm are two different components. A full description of the notion of component is in Chapter 2, and pictographic examples are provided in section 2.6.

A full explanation of components and how they were allocated is presented in Chapter 2. The parameters of iconicity have been detailed in section 1.2. These are taken directly from Poggi (2008) and have been used in studies by Merola (2007). The four main parameters are shape, location, agent action to referent, and referent’s typical action. The four subsidiary parameters are the cause-effect relationship, the
object-function relationship, the class-prototype relationship, and the opposition-negation relationship.
2 Methodology

To test the research hypotheses and the research question, the experiment required participants to perform a task similar to the game of charades. Ten pairs of participants were asked to convey 120 concepts to each other (60 concepts each) through gesture alone. These concepts were made up of 20 individual concept tokens with 6 iterations for each concept. Participants took it in turns to be gesturers and guessers. After gesturing 10 concepts, participants swapped roles. At the completion of the experiment, a short interview was conducted in which each pair had to identify their rationale for some of the gestures they chose to perform. The variables for this experiment are listed below, in their relevant subgrouping (dependent and independent).

2.1 Variables and Experimental Design

There were four variables for this experiment. For hypothesis 1 simplicity was measured by the number of components: fewer components indicated a greater level of simplicity. The dependent variable for hypothesis 1 was the number of components used to express a single meaning. The independent variable was the number of iterations per concept. For hypothesis 2 the selection of iconic parameters to represent different concept types was measured. The dependent variable for hypothesis 2 was the type of iconicity used in gesture (shape, location, agent’s action, and referent’s typical action). The independent variable was the concept type of each experimental concept (action, artefact, natural object, and animate).

The types of iconicity used in gesture will follow the parameters defined in Poggi (2008). To summarize, the four main types of iconicity are referent shape,
referent location (non-deictic), referent’s typical actions, and agent action on referent. There are four inferential relationship types, which are used to depict abstract concepts (see section 1.2). As seen in Merola (2007), the types of iconicity employed by the gesturer will have a connection to the types of concept used. The four types of concepts in this experiment are the same as in Merola (2007): action, artefact, natural, and animate. Action words are verbs, which means that the words refer to actions that humans perform. Artefacts are items created by humans, and they can be interacted with in some way. Natural objects are referents found in nature, detailed clearly under the term 'natural kind' (Quine, 1969). Animate objects are living beings that move. The full list of concepts is given in Table 2. These concepts were based on the word list in Merola (2007).

Table 2

Experimental Concept List

<table>
<thead>
<tr>
<th>Action</th>
<th>Artefact</th>
<th>Natural</th>
<th>Animate</th>
</tr>
</thead>
<tbody>
<tr>
<td>to think</td>
<td>bag</td>
<td>mountain</td>
<td>fish</td>
</tr>
<tr>
<td>to crush</td>
<td>door</td>
<td>sun</td>
<td>caterpillar</td>
</tr>
<tr>
<td>to find</td>
<td>wheel</td>
<td>tree</td>
<td>mosquito</td>
</tr>
<tr>
<td>to tie</td>
<td>chair</td>
<td>cloud</td>
<td>child</td>
</tr>
<tr>
<td>to swing</td>
<td>boat</td>
<td>river</td>
<td>dog</td>
</tr>
</tbody>
</table>

A pilot study was conducted using 68 different concept tokens with no iterations. From the 68 concept tokens 20 were selected. Concepts were excluded from the final concept list based on how difficult the participants in the pilot study found it to convey the meaning, whether participants used numeric gesture (such as using their fingers to display number), and whether the word used for the concept had multiple possible meanings. Concepts were also excluded if participants included
obscene gestures or pre-existing cultural gestures. Some concepts proved harder to represent in the pilot study, for example valley was difficult for both pilot study pairs, while participants did not struggle as much with mountain.

The number of components included within gesture will be used as a measure of simplicity. The number of components is expected to reduce across subsequent iterations, firstly due to mutual understanding being established and secondly due to participants’ attempting to simplify gestures. To reiterate what was mentioned in section 1.4, a component is a compact bundle of gestures that relate to the same semantic meaning. A component was subjectively decided when it was apparent that the participant was gesturing a new meaning. For example pinching fingers together to depict “small”, and ‘crawling’ a finger along the forearm are two different components. Components are also apparent during single-concept iterations. Though components were apparent, the precise boundary of each component was unclear. Żywiczyński et al. (2016) stated that pantomimic gesture has “no self-apparent onsets and terminations in the stream of movement” (p. 8). There is currently no established method to detail the onsets and terminations of a component. Component boundaries were thus subjectively decided. If a gestural component could switch places with another, or if it was able to be excluded, then it was clear that the bundled meaning was a self-contained constituent.

The number of iterations for each concept was consistent – iterations were not manipulated within the experiment. However, all iterations were set deliberately to examine how iconic gestures became abstracted. Because of this, iterations are

---

3 The establishment of mutual understanding is also known as ‘grounding’. The terms are used interchangeably in this research.
counted as independent variables. Each concept was repeated six times in total, three times by each participant. The full list of every concept, their iterations, and the order they appeared in the experiment can be seen in Table 3.

The concepts were presented in random order. The first 20 concepts were randomized separately from the subsequent 100 to allow each participant to establish gesture for 10 concepts each. The remaining 100 concepts were then randomized by the same method. The concept list was manually adjusted to remove any immediate doubles, e.g., cloud followed by cloud, as well as triple occurrences of a concept within one set. This prevented any recognition of the pattern in which concepts would occur. Table 3 lists all 120 concept tokens in the order they were presented. Each column shows the ten concepts that each participant gestured, and is read from top to bottom, left to right.

Table 3

<table>
<thead>
<tr>
<th></th>
<th>Concept Tokens 1-60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A-1</td>
</tr>
<tr>
<td>1</td>
<td>to think</td>
</tr>
<tr>
<td>2</td>
<td>caterpillar</td>
</tr>
<tr>
<td>3</td>
<td>door</td>
</tr>
<tr>
<td>4</td>
<td>fish</td>
</tr>
<tr>
<td>5</td>
<td>tree</td>
</tr>
<tr>
<td>6</td>
<td>bag</td>
</tr>
<tr>
<td>7</td>
<td>mountain</td>
</tr>
<tr>
<td>8</td>
<td>to find</td>
</tr>
<tr>
<td>9</td>
<td>to swing</td>
</tr>
<tr>
<td>10</td>
<td>chair</td>
</tr>
</tbody>
</table>
2.2 Participants

Twenty speakers of English participated. The majority of the participants were first language speakers of English, and two participants were second language speakers of English. Participants were between 18 and 35 years of age. Most participants were students at the University of Otago and the others were residents of Dunedin. Relevant information on participants is included in Table 4. The participant’s code was created based on the order in which their pair participated (1-10), followed by the position they filled (A or B). All participants completed a demographic form prior to commencing the experiment. Demographic data collected included: their age, gender identity, country of birth, ethnicity, languages spoken (and the length of time speaking it), places they have lived other than New Zealand (if anywhere), whether they knew their paired partner, and whether they had any physical difficulty with gesture.
As different languages have different culturally-based gestures, participants were asked to state any languages they knew to help account for any potential gestural anomalies. Participants ranged from speaking one language (English) to four languages. The majority of participants had been learning languages at an educational institute (either secondary or tertiary) and were therefore unlikely to have adopted culturally-based gestures associated with these languages. Eight participants were raised in another English-speaking country (America, Britain, or Australia), and two participants were raised in non-English speaking environments (Germany and Lithuania). All ten of these participants have been exposed to a different style of cultural gesture, especially the two participants from non-English speaking environments. With most participants having low levels of proficiency in their other languages, it is assumed for the most part that additional languages will not be a confounding factor. Participants were asked if they knew their paired partner to rationalize any pre-existing gestures that might exist between people who interact on a regular basis.

Table 4

Participant Information

<table>
<thead>
<tr>
<th>Code</th>
<th>Sex</th>
<th>Geographic Origin</th>
<th>Other languages</th>
<th>Know partner?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>M</td>
<td>UK</td>
<td>SPA (9yrs)</td>
<td>N</td>
</tr>
<tr>
<td>1B</td>
<td>F</td>
<td>NZ</td>
<td>JAP (2yrs), SPA (6mnth)</td>
<td>N</td>
</tr>
<tr>
<td>2A</td>
<td>F</td>
<td>NZ</td>
<td>SPA (3yrs)</td>
<td>N</td>
</tr>
<tr>
<td>2B</td>
<td>F</td>
<td>USA</td>
<td>SPA (5yrs)</td>
<td>N</td>
</tr>
<tr>
<td>3A</td>
<td>F</td>
<td>NZ</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>3B</td>
<td>M</td>
<td>NZ</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>4A</td>
<td>F</td>
<td>USA</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>4B</td>
<td>M</td>
<td>NZ</td>
<td>JAP (3yrs)</td>
<td>N</td>
</tr>
<tr>
<td>5A</td>
<td>F</td>
<td>AUS</td>
<td>JAP (5yrs)</td>
<td>N</td>
</tr>
<tr>
<td>5B</td>
<td>M</td>
<td>NZ</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>6A</td>
<td>F</td>
<td>LT</td>
<td>LT (26yrs), GER (14yrs), NOR (5yrs), ENG (19yrs)</td>
<td>Y</td>
</tr>
<tr>
<td>6B</td>
<td>M</td>
<td>USA</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>7A</td>
<td>F</td>
<td>USA</td>
<td>FRE (10yrs), ASL (5yrs)</td>
<td>N</td>
</tr>
<tr>
<td>7B</td>
<td>F</td>
<td>NZ</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>8A</td>
<td>M</td>
<td>NZ</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>8B</td>
<td>M</td>
<td>NZ</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>9A</td>
<td>M</td>
<td>UK</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>9B</td>
<td>F</td>
<td>NZ</td>
<td>FRE (4yrs), GER (4yrs), SPA (2yrs)</td>
<td>N</td>
</tr>
<tr>
<td>10A</td>
<td>F</td>
<td>GER</td>
<td>GER (24yrs), FRE (10yrs), MAN (2yrs), ENG (12yrs)</td>
<td>N</td>
</tr>
<tr>
<td>10B</td>
<td>M</td>
<td>UK</td>
<td>FAR (7yrs)</td>
<td>N</td>
</tr>
</tbody>
</table>

**KEY:**

FRE = French  
GER = German  
SPA = Spanish  
LT = Lithuanian  
UK = United Kingdom

FAR = Farsi  
MAN = Mandarin  
JAP = Japanese  
NOR = Norwegian  
ASL = American Sign Lang.

**NOTE:**

All participants are fluent in English. Two participants have English listed as a language only because it is their second language.
One participant had been learning and using ASL for 5 years. Their knowledge of a gestural language was mitigated by pairing them with a partner that did not know ASL. This is because their paired partner had no previous experience with a gestural language, meaning that the participant experienced with ASL was unable to benefit from previously-existing cultural forms of gesture.

Participants were not at risk during the experimental process. This study obtained ethical approval from the University of Otago ethics committee.

2.3 Experiment Set-Up and Equipment

The experiment was run in the same room within the University of Otago Dunedin campus. Two cameras were used to record the procedure, which were left at set positions throughout the duration of all ten experiments. The camera used to record gesturer movement was a Panasonic HC-V770, and the camera used to record the guesser was a Canon EOS 700D. The 120 experiment cards were printed on white paper and cut to 4.5cm height and 9cm width. Size 36 Calibri text was used for ease of reading. On one side of the card the concept token was printed, and on the other side the position of the concept within the set was listed. Figure 1 shows the card layout for the concept bag. It was the sixth concept to appear in the first set of concepts for participant A. There were 12 sets of 10 concepts each, 6 sets for each participant.
The room was laid out as seen in Figure 2. The guesser was seated on a chair facing the gesturer, while the gesturer was allotted a space in which they could freely move. This space was marked by tape on the ground, and participants were instructed to avoid crossing the tape during their gestures due to camera limitations. The 12 sets of cards were laid out on the seat within the gesturers ‘zone’, allowing participants easy access to the cards. In front of the chair, there was a box for the gesturer to discard their card when they had finished with it. The cameras were in set positions on either corner of the room, angled to record each participant. The researcher was seated near the guesser camera. Any posters or wall hangings were removed so that walls were bare prior to the beginning of the experiments.
2.4 Procedure

The experiment was conducted as follows:

First, the participants were greeted and given the “Information for Participants” sheet, “Consent Form”, and the participant demographic form. After reading, signing, and filling out the respective forms, the researcher explained the experiment. The explanation followed the outline in Figure 3.

Hello and welcome to the gesture experiment!

First thank you for being here and participating. I really do appreciate it.

Second Please, read through the information sheet. The most important points are:

- you agree to be recorded
- you understand pictures from the filming can be used in the thesis
- you’re welcome to have a copy of the thesis when it is done!

Third Once the cameras are rolling, I will say “okay, start!”. From this point, the gesturer is to pick up a card, read it, put it back down, and then gesture it. The gesturer can hold on to the card for a short while as they consider how they might gesture it. Please try to stay in the tape-marked area, otherwise I might miss some of your movements!

The guesser has two guesses – Be aware that you want to give one guess at a time! The gesturer needs a chance to “recast” the gesture.

If the meaning is not conveyed – Don’t worry! You have to just move on. There will be another chance to do that word.

After 10 words, gesturer and guesser switch – This is clearly marked on the card sets!

Please avoid pointing at any objects in the room that bear resemblance to the word on the card. Also, please avoid making noises while gesturing where possible.

Once the experiment starts, we will continue until all words have been completed.

After the experiment finishes, I will ask a few questions about things I found interesting and to find out how you felt about it. This will be recorded, but only so that I don’t forget or miss anything.

Figure 3. Instruction outlines from the experiment.
After the instructions, a test period was offered. Each participant took a test card, one with the word *burger* and the other with the word *bath*. This trial was offered to allow participants a chance to actively understand what was expected of them, and to allow for any questions they might have. The test phase also allowed for early allocation of ‘A’ role and ‘B’ role, which meant that the gesturer and guesser roles were easily established. After the test phase, participants continued with the experiment, moving from A1 to B1, right through until they reached A6 and B6. The researcher had minimal interaction with participants during the running of the experiment, with the exceptions of answering any questions or expressing that a guessed concept was close enough to suffice (for example, *backpack* was an acceptable near guess for *bag*). The average duration of each experiment was 20 minutes. Following the completion of the experiment, participants were interviewed informally. During this time, the researcher asked questions about the types of gestures used, as well as questions about participant rationale regarding the gestures performed. Finally, participants were given a $5 note and a small gift as a show of appreciation for the time they had volunteered.

2.5 Analysis

The video files from both cameras were transferred to a computer. Using Windows Movie Maker the video files were spliced into smaller files, each file containing only the gestures for a concept. In total, each pair had 120 individual

---

4 Three different participants had initially guessed *backpack* for *bag*. Once both participants in the respective pairs had read the experimental card the guess was consistently *bag*. Other instances of near guesses were uncommon.
gesture video files and there were 1,200 video files overall. These files were converted into an mp4 format.

Each concept was analyzed first by how many components were involved, and then by how many different parameters of iconicity were included within each component. The parameters of iconicity have been detailed in the previous chapter. These are taken directly from Poggi (2008) and have been used in studies by Merola (2007). The four main parameters are shape, location, agent action to referent, and referent’s typical action. In the description provided in section 1.2, the description of shape does not define whether aspects such as size are included. In this research, size is considered as an aspect of shape. In addition to this, any depiction of a singular body part, for example depicting the proboscis of a mosquito, is also considered as an aspect of shape. After all gestures were coded in respect to their components and levels of iconicity present, the data were run through R-Studio (R Core Team, 2013). The data were analyzed with a combination of descriptive statistics and multi-level regression.

During the data cleaning process six pieces of data were removed and three were altered. The six pieces that were removed used symbolic forms and lexical gestures. The symbolic forms were tracing the cartoon image of the sun\(^5\), and using cultural dance moves. Participants used lexical gestures to negate meaning, to display an ordinal number, and to inflect future events. All of these symbolic and lexical gestures are discussed further in Chapter 4. The three pieces that were altered were due to abnormally high component counts in either iteration two or three. Two of these were because of the guesser did not guess correctly in iteration three, even

\(^5\) Two instances of this were removed.
after two previous successful guesses. Sequentially from the first iteration to the last, the component count of the first instance was 4, 5, 10, 2, 4, and 1. The component count of the third repetition was changed to 5, which was the component count of the previous successful guess. The component count of the second instance was 3, 3, 9, 1, 1, and 4. The component count of the third iteration was also changed to match the previous successful guess, which was 3. The third piece of data that was altered had component counts of 4, 13, 3, 1, 1, and 1. During the second iteration, the gesturer stood at a different angle from the one in which the other participant had stood in the first iteration. The change of angle was the participant’s stated cause for not understanding quicker, as they attested during the experiment after guessing correctly. The component count of the second iteration was changed from 13 to 4 to match the previous successful iteration.

2.6 Example: *mosquito*

This section provides a full example of how components were determined. The example is participant 4B gesturing the concept *mosquito*. The following 50 images are screenshots from 9 seconds of video. Each image is approximately 0.16s apart, which means that every 6 images depict gestures that occur over the course of one second. Underneath each line of three images a description of the gestures is given, along with a description of which iconic parameters are being used.
Model (1) shows the participant's final moment of thinking before initiating the first component. Participant 4B is not engaged in any gesture other than those relating to thinking. Model (2) shows the first motions in the component, which flow directly from the previous gesture of thinking. In this component the participant depicts 'small', which is an aspect of size. This matches the iconic parameter of shape. The beginning of this gesture is the onset of the first component. Models (3), (4) and (5) are continuations of the depiction of 'size'. Model (6) is the first gestural motion of the second component, which is the depiction of wings flapping. Again, the participant fluidly transitions from one gesture to the next, overlapping component boundaries.
The onset of the second component is also the termination of the first component. There is no pause between the two; the participant fluidly transitions between finishing the first component and commencing the second component. Models (7) to (14) all show the depiction of wings flapping, which is a continuation of the second component. The action of flapping is classified as a referent’s typical action.

Figure 6. Model (7), (8), & (9)

Figure 7. Model (10), (11), & (12)
Model (15) shows the participant finishing the second component, which is the termination of the component. Models (16), (17), and (18) show intermediary gestures. The participant is moving his hands towards the onset of the next component. The onset could be said to be at any point during the intermediary stage. The onset is deemed to be at Model (19), where participant 4B looks at his paired partner. This third component is the participant gesturing an object which is both long (indicated through moving the right hand away from the left hand in a straight line) and thin (indicated through the ‘pinched’ index fingers to the thumbs). This component continues across Models (20) and (21). The depicted object is the proboscis of a mosquito, which is apparent from the place of the participant’s hands relative to his face. As mentioned earlier, actions which depict a body part are
classified as shape. This component, a depiction of a proboscis, is classified as an iconic parameter of shape. It could also be classified as a typical action of a referent, but as the participant does not mime using the proboscis to suck blood, it is instead classified as shape.

![Figure 10. Model (19), (20), & (21)](image)

![Figure 11. Model (22), (23), & (24)](image)

Model (22) is the last frame where the participant uses gestures to depict a proboscis, and therefore it is classified as the termination of the third component. Model (23) is the onset of the fourth component, which is a depiction of a mosquito biting the participant’s arm. This action is classified as a typical action of a referent. Models (24) to (28) show the fourth component being performed, and Model (29) is
the termination of the component as it is the final gestural motion that depicts the bite.

Figure 12. Model (25), (26), & (27)

Figure 13. Model (28), (29), & (30)

Figure 14. Model (31), (32), & (33)
Models (30) to (33) show the participant shaking his head, indicating ‘no’. This is because the guesser had made a guess of ‘hummingbird’. This type of communicative gesture was not counted as an iconic parameter, and the component was not documented. Models (34) to (38) show the fifth component being gestured. The component is the depiction of ‘small’ again. Model (34) is the onset and Model (38) is the termination. Note how quickly the participant shifted from indicating ‘no’ in Model (33) to beginning to depict size in Model (34). This example shows a clear onset point.

Figure 15. Model (34), (35), & (36)

Figure 16. Model (37), (38), & (39)
Figure 17. Model (40), (41), & (42)

Model (39) is the onset for the sixth component. It is another example of a clear onset. Models (40) and (41) are continuations of the component, which depicts a mosquito biting the participant, the same as in the fourth component. It is a depiction of a typical action of the referent. Model (42) is the termination of the component.

Figure 18. Model (43), (44), & (45)

The onset for the seventh component is Model (43) and the termination is Model (47). The participant is rubbing his arm and depicting pain on his face, which is an agent’s action. The typical action of the referent and the action of the agent are clearly contained within separate components, and not mixed at any point. The
termination of the component is indicated by the look of pain suddenly disappearing from the participants face.

*Figure 19. Model (46), (47), & (48)*

Models (48) and (49) show the participant listening to his paired partner making their second guess of ‘mosquito’. Model (50) shows participant 4B signaling that the guesser was correct by nodding and pointing.

*Figure 20. Model (49), & (50)*

These 50 models showed the seven different components participant 4B used to convey *mosquito* on the first iteration of the concept. In the description of the examples, both the onset and the termination of each component were stated. The boundaries of some components were immediately next to each other, as one gesture
flowed on to the next one. While the general outlines of the boundaries were apparent, the precise time of onset is unclear. Further examples are in sections 3.3.1 (Example: bag) and 3.3.2 (Example: mountain). While those examples do not point out the onset and termination of a component, there are gestures that have clear boundaries which others do not have. The imprecise boundaries of components does not impact on the existence of a component, so that analyzing component use is nonetheless possible. Spatial and temporal analysis of gestures requires a precise onset and termination location, which is discussed further in Chapter 4.
3 Results

The current research sought to confirm two different hypotheses as well as a central research question. Hypothesis 1 states that participants will simplify their iconic gestural components following the establishment of mutual understanding. Hypothesis 2 states that participants will choose to represent action word types by agent action, artefacts by agent action, natural objects by shape, and animates by the typical action of a referent. The main research question for this thesis is “how do iconic manual gestures become simplified?” Each of two hypotheses requires different types of analysis to be answered, and the research question requires the analyses from these hypotheses along with discourse about the question to be answered.

As detailed in Chapter 2, the experiment for this study was involved elements similar to those in the game of charades. Participants were organised into pairs and were given 120 concept tokens, comprised of 6 iterations of 20 concepts. Participants were allocated to a ‘gesturer’ or ‘guesser’ role, which switched after ten concepts. The gesturer was asked to convey each concept through gesture alone, and the guesser was asked to guess what the gesturer was conveying. Throughout this process, the participants were video recorded for video analysis. Details of the analyses can be found in sections 2.5, 3.1, and 3.2.

This chapter is organised as follows: the first hypothesis is restated and the supporting data are presented in section 3.1, the second hypothesis is restated and the supporting data are presented in section 3.2, and then the central research question is restated and considered in light of the experiment results in section 3.3. There are two examples of simplification within the experiment: one for the concept
of bag and one for the concept of mountain. The research question will be explored in more depth in Chapter 4.

### 3.1 Hypothesis 1

Hypothesis 1 states that participants will initially construct iconic forms of gesture. Iterations following the establishment of mutual understanding will lead to simplification of the initial iconic form. Hypothesis 1 is comprised of three beliefs: 1) participants will gesturally construct the first iteration of a concept with a heavy reliance on iconic forms, 2) when the researcher compares the first iteration of a concept with later iterations of the same concept, iconicity will have reduced, and 3) iterations after the successful conveyance of a concept will be simplified. As discussed in section 2.1, hypothesis 1 will be confirmed by using the variables of number of components (dependent) and the number of iterations for a concept (independent). Components are bundles of smaller gestural parts that comprise a semantic whole.

While a component is not necessarily iconic, participants rarely used non-iconic gestures. 2,508 components were documented as iconic, but only five components were documented as non-iconic. These five will be discussed in Chapter 4. However it important to note that components were primarily iconic. After removing the non-iconic components, all remaining components were listed as iconic. A decrease in the number of components across six iterations can be considered as a decrease in iconic forms; in other words, the set of components should simplify across iterations.
Figure 21: All participants’ average number of components per iteration

Table 5

<table>
<thead>
<tr>
<th>Rep1</th>
<th>Rep2</th>
<th>Rep3</th>
<th>Rep4</th>
<th>Rep5</th>
<th>Rep6</th>
</tr>
</thead>
<tbody>
<tr>
<td>sd</td>
<td>3.8</td>
<td>2.5</td>
<td>1.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Figure 21 shows a larger average number of components for the first iteration in comparison with iterations 2 to 6. This could be due to participants establishing mutual understanding, which will be discussed further in Chapter 4. There is a decrease in average number of components from iteration 2 through to iteration 4, after which components remain stable. The graph above is drawn from data which includes both successful and unsuccessful attempts at conveying the concept. Figure 22 is based on successful attempts only \(^6\). The same general trend is observed;

\(^6\) Note that unsuccessful attempts were most prevalent in the first iteration (42 non-successes). The second iteration had 6 unsuccessful attempts.
iteration one has the highest average number of components, iteration two has fewer, iteration three has fewer still, and then the following iterations plateau.

![Average Components per Iteration (Success Only)](image)

*Figure 22: All participants’ average number of components per iteration, excluding unsuccessful attempts*

**Table 6**

*Standard Deviations of Components per Iteration (Success Only)*

<table>
<thead>
<tr>
<th></th>
<th>Rep1</th>
<th>Rep2</th>
<th>Rep3</th>
<th>Rep4</th>
<th>Rep5</th>
<th>Rep6</th>
</tr>
</thead>
<tbody>
<tr>
<td>sd</td>
<td>4</td>
<td>1.6</td>
<td>1</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

To find any abnormalities in the data from specific participant pairs, similar graphs were created that considered a pair at a time. Figure 23 is the amalgamated version of these ten graphs. This graph was created to show all pairs simultaneously.
Figure 23: Average components of all pairs

Figure 24: Average components of all pairs with unsuccessful attempts removed
Figure 24 shows the average number of components from successful attempts only. In comparison with figure 23, the average number of components is higher. This is especially true of pair 9's iteration one average, which moves from 6 in Figure 23 up to 11 in Figure 247. The number of components in successful attempts was varied. Sometimes the guesser would guess correctly after three components, and sometimes the guesser would wait until the gesturer had completed a much larger set of gestures (e.g., pair 9’s attempt of chair took 20 components before a successful guess). Performed gestures that led to an unsuccessful guess ranged from having one component through to have 15 components, with an average of five components.

Figure 24 also shows that there is very little difference overall to the components after iteration 2. Both Figure 23 and Figure 24 show that the majority of pairs consistently used a similar number of components per concept. With the exception of iteration one, all of the averaged components have a similar count.

7 As pair 9 have higher sets of components compared to other pairs, an averaged total of components per iteration was also done which excluded pair 9. The averaged total (unclean) data were 3.9, 2.3, 1.9, 1.5, 1.5, and 1.5. Compared with figure 21, there is little difference. Pair 9’s high component count was deemed unnecessary to remove due to this minimal difference in averages.
Figure 25: Average components of all four concept types: artefact, action, animate, and natural objects

Figure 25 separates the average components by concept type (artefact, action, animate, and natural object). The same general trend from the previous graphs is observed again: iteration one has a higher average component count, which then drops with iteration two, and plateaus after iteration three. The only exception to this is iteration three of ‘natural objects’.

Multi-level linear regression was used to test the components present per iteration for iterations 1 to 6, 2 to 6, and 3 to 6. A regression model was constructed to predict Component Count with fixed effects of Iteration, and a random effect of Pair. All statistics were implemented using the R statistical software (R Core Team, 2013) with lme4 (Bates, Maechler, Bolker, & Walker, 2015) and lmertest (Kuznetsova, Brockhoff, & Christensen, 2017). The values show that the components of iterations 1 to 6 have a significant chance of decreasing ($t=-11.86$, $p<0.001$), the components of iterations 2 to 6 also have significant chance of decreasing ($t=-2.73$, $
$p<=0.001$), and the components of iterations 3 to 6 do not have a significant chance of decreasing ($t=7.08, p=0.48$). These values are shown in Table 7.

Table 7

**Multi-Level Linear Regression Examining the Effect of Iteration on Component Count**

<table>
<thead>
<tr>
<th>All iterations (1-6)</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-11.858</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Only iterations 2-6</td>
<td>-2.733</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Only iterations 3-6</td>
<td>7.08</td>
<td>0.48</td>
</tr>
</tbody>
</table>

As stated, the research hypothesis claimed: 1) participants will gesturally construct the first iteration of a concept with a heavy reliance on iconic forms, 2) when the first iteration is compared with later iterations of the same concept, the level of iconicity will decrease, and 3) iterations following a successful attempt will be simplified. All three of these claims appear to be confirmed in figures 21-25.

Participants did have a heavy reliance on iconic gesture. The number of components does decrease from iteration one through to six. There is a simplification of the components over time, which is an early stage of abstraction. Finally, iterations that followed a successful attempt were simplified (in that they were comprised of fewer components).

### 3.2 Hypothesis 2

Hypothesis 2 predicts that participants will gesture certain types of concept classes according to iconic parameters as set by Poggi (2008).
1. the REFERENT'S SHAPE: to mean “mountain” you can outline a conic shape, for “brioche” your hand forms a rounded shape

2. the REFERENT'S LOCATION: to mean “sign” you draw a rectangular shape up over you (sic) head; for “mountain” you can point far away

3. the REFERENT'S TYPICAL ACTIONS: to mean “bird” you move hands like wings for “lightning” you depict its zigzag trajectory

4. an AGENT'S ACTION with the referent: to convey “mountain” you pretend climbing or skiing

(Poggi, 2008, p. 52)

These four parameters were observed in a study by Merola (2007) on iconic gestures used by school teachers while instructing new vocabulary. Table 1, on page 21, shows the results. While Merola made no overt predictions that a certain concept type would cause participants to choose different iconic parameters, there is a trend in Merola’s data that should be repeatable. This trend is that participants will choose to represent action word types by agent action, artefacts by agent action, natural objects by shape, and animates by referent typical action.
Table 8

Iconic Parameters per Concept Type (All Iterations) a

<table>
<thead>
<tr>
<th>Concept Type</th>
<th>Shape</th>
<th>Location8</th>
<th>Referent Action</th>
<th>Agent Action</th>
<th>TOT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n.</td>
<td>%</td>
<td>n.</td>
<td>%</td>
<td>n.</td>
</tr>
<tr>
<td>Action</td>
<td>38</td>
<td>7</td>
<td>15</td>
<td>3</td>
<td>510</td>
</tr>
<tr>
<td>Artefact</td>
<td>201</td>
<td>27</td>
<td>69</td>
<td>10</td>
<td>436</td>
</tr>
<tr>
<td>Natural Object</td>
<td>331</td>
<td>46</td>
<td>171</td>
<td>23</td>
<td>155</td>
</tr>
<tr>
<td>Animate</td>
<td>93</td>
<td>17</td>
<td>350</td>
<td>66</td>
<td>91</td>
</tr>
<tr>
<td>TOT</td>
<td>663</td>
<td>-</td>
<td>605</td>
<td>-</td>
<td>1192</td>
</tr>
</tbody>
</table>

a The highlighted cells indicate the highest percentage of an iconic parameters use for each concept type.

Table 8 shows that participants chose to represent an action by using an agent’s action 510 times (90%). Artefacts were represented by an agent’s action 436 times (60%). Natural objects were represented by shape 331 times (46%). Animates were represented 350 times (66%) by a typical action of the referent.

These results are very similar to the results from Merola (2007). The notable difference is that participants represented natural objects by an agent’s action and by a referent’s typical action, while participants in Merola’s study did not choose to represent natural objects by either of these parameters. This could be due to this study having a larger sample size9, or it could be due to a difference of representation location compared to Merola’s study.

---

8 Location was separated from deictic pointing. It was most commonly used for the word ‘bag’, where participants indicated the location of the bag (on their back) while depicting the bag’s shape. Location was also used in natural objects, primarily for the word ‘sun’. Participants indicated the location of the sun while depicting the shape.

9 Merola had four teacher participants who were videotaped while teaching 24 different words.
between pantomimic gesture and paralinguistic gesture. This will be discussed in Chapter 4.

Table 8 is a summation of all gestures used across all iterations. While it is useful to consider a total sum of the gestures, participants often chose a variety of iconic parameters to represent a concept during the first iteration. This was useful, if not essential, to the development of mutual understanding. However, the variety of gestures used in early iterations can impact on the clarity of results. Table 9 excludes all iterations except for iteration 6. By iteration 6, all participants had decided on a fixed representation of the concept.

Table 9

.Iconic Parameters per Concept Type (Iteration 6 Only)\(^a\)

<table>
<thead>
<tr>
<th>Concept Type</th>
<th>Shape</th>
<th>Location</th>
<th>Referent Action</th>
<th>Agent Action</th>
<th>TOT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n.</td>
<td>%</td>
<td>n.</td>
<td>%</td>
<td>n.</td>
</tr>
<tr>
<td>Action</td>
<td>4</td>
<td>6</td>
<td>-</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Artefact</td>
<td>21</td>
<td>23</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Natural Object</td>
<td>42</td>
<td>54</td>
<td>5</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>Animate</td>
<td>11</td>
<td>16</td>
<td>-</td>
<td>49</td>
<td>70</td>
</tr>
<tr>
<td>TOT</td>
<td>78</td>
<td>-</td>
<td>7</td>
<td>79</td>
<td>-</td>
</tr>
</tbody>
</table>

\(^a\) The highlighted cells indicate the highest percentage of an iconic parameter’s use for each concept type.

Table 9 shows the same patterns as Table 8. This table shows that, at iteration 6, participants chose to represent an action by using an agent’s action 59 times (91%). Artefacts were represented by an agent’s action 59 times (66%). Natural
objects were represented by shape 42 times (54%). Animates were represented by a referent’s typical action 49 times (70%).

The close similarity in percentages between Table 8 and Table 9 shows that there is little difference between data from an amalgamation of all iterations and data from the final iteration. The similarities between Merola’s data (Table 1, page 21) and the data from the current research (Table 8 and Table 9) show that concept types (natural objects, animates, artefacts, and actions) are preferably represented by certain iconic parameters. In particular, participants chose to represent both actions and artefacts by an agent’s action, natural objects by their iconic shape, and animates by their typical actions (referent’s typical action).

Multilevel logistic regression was used to test the use of an iconic parameter against some concept types compared to other concept types. A regression model was constructed to predict Iconic Parameter with fixed effects of Iteration and Concept Type, and a random effect of Pair. It was found that shape is used predominantly with natural objects when compared with actions (Wald’s Z=-15.103, p=< 0.001, Table 10), artefacts (Wald’s Z=-9.670, p=< 0.001, Table 10), and animates (Wald’s Z=-12.758, p=< 0.001, Table 10). Agent action was used predominantly with artefacts compared with animates (Wald’s Z=-14.079, p=< 0.001, Table 11) and natural objects (Wald’s Z=13.423, p=< 0.001, Table 11). Agent action was also used predominantly with actions compared with animates (Wald’s Z=-19.796, p=< 0.001, Table 11) and natural objects (Wald’s Z=-19.796, p=< 0.001, Table 11). A referent’s typical action was used predominantly with animates when compared with artefacts (Wald’s Z=-17.887, p=< 0.001, Table 12), natural objects (Wald’s Z=-13.121, p=< 0.001, Table 12), and actions (Wald’s Z=-15.069, p=< 0.001, Table 12). Location was
used for artefacts or natural objects only and was slightly more prevalent in natural objects than artefacts ($Wald's Z=6.533, p< 0.001$, Table 13).

**Table 10**
*Chance of ‘Shape’ Parameter Being Selected for Different Concept Types Using glmer*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Concept Type 1</th>
<th>Concept Type 2</th>
<th>Wald’s Z</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>Animate</td>
<td>Artefact</td>
<td>4.940</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>12.757</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>-5.039</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Artefact</td>
<td>Animate</td>
<td></td>
<td>-4.941</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>9.670</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>-9.208</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Natural Object</td>
<td>Artefact</td>
<td></td>
<td>-9.670</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Animate</td>
<td>-12.758</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>-15.103</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Action</td>
<td>Arтеfact</td>
<td></td>
<td>9.208</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Animate</td>
<td>5.039</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>15.103</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Table 11**
*Chance of ‘Location’ Being Selected for Different Concept Types Using glmer*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Concept Type 1</th>
<th>Concept Type 2</th>
<th>Wald’s Z</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent Action</td>
<td>Animate</td>
<td>Artefact</td>
<td>14.079</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>1.674</td>
<td>0.0941</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>19.796</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Artefact</td>
<td>Animate</td>
<td></td>
<td>-14.079</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>-13.423</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>9.802</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Natural Object</td>
<td>Artefact</td>
<td></td>
<td>13.423</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Animate</td>
<td>-1.674</td>
<td>0.0941</td>
</tr>
<tr>
<td>Parameter</td>
<td>Concept Type 1</td>
<td>Concept Type 2</td>
<td>Wald’s Z</td>
<td>p value</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Referent Action</td>
<td>Animate</td>
<td>Artefact</td>
<td>-17.887</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>-13.121</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>-15.069</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Action</td>
<td>Artefact</td>
<td>Animate</td>
<td>17.887</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>6.518</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>-4.464</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Artefact</td>
<td>Animate</td>
<td>Artefact</td>
<td>-6.518</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Natural Object</td>
<td>13.120</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Action</td>
<td>-8.484</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Natural Object</td>
<td>Artefact</td>
<td>Action</td>
<td>4.464</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Action</td>
<td>Artefact</td>
<td>Natural Object</td>
<td>15.068</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Artefact</td>
<td>8.484</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 12
Chance of ‘Referent Action’ Being Selected for Different Concept Types Using glmer

Table 13
Chance of ‘Location’ Being Selected for Different Concept Types Using glmer
3.3 Research Question

The central research question for this thesis is:

“How do iconic manual gestures become simplified?”

By coding gestures into component parts it was possible to observe simplifications occurring to components. This is apparent in the analysis of hypothesis one, which showed that the average number of components decreased across the course of four iterations. The final two iterations were similar enough to iteration four to be considered to have plateaued. A reduction in the number of components used to convey the concept means that the number of gestures has reduced. Other simplifications were classified, for example participants were much more careful and considered with their gestures during the first iteration, but far less so during the sixth iteration. Specifically, it was noticed that participants performed the gesture more quickly and within a smaller gestural space. Ideally, a researcher could use motion sensor technology to precisely map both the temporal aspects of gesture and the spatial aspects. This would allow for speculations on the simplification process. While the current research does not employ motion sensor technology, the simplifications observed can inform an outline of simplification. The three general stages of simplification that were noticed are:

1. The gesturer chose to use either one gesture or a series of iconic gestures until the guesser correctly guessed the concept. This was either a single gesture or a long set of components, depending on how well the gesturer depicted the signified object, how complex the series of actions were, and how well the guesser interpreted the gesture/s.
2 - On the next iteration, the gesturer chose components perceived the most salient and repeated them. At this point, most participants were consciously attempting to keep the gesture as similar as possible to those in the first iteration.\(^\text{10}\)

3 - As both participants became habituated to the experiment, the gesturer used fewer components with the expectation that the guesser would “get it” after the provided gestures. With both the guesser and the gesturer attempting to achieve the task more quickly, gestures became faster and less precise. The number of components used remained stable or had slight variations (variation came from the participants repeating gestures, participant difference in conceptualization, as well as from participants not being able to complete the entire set of components because the guesser responded too quickly).

These three stages are shown in the following two examples. The first example includes all six iterations of pair 9 gesturing \textit{bag} (Figures 26 – 43). The second example includes all six iterations of pair 5 gesturing \textit{mountain} (Figures 44 – 55). In both examples gestures reduce the number of components they use across the course of the iterations. These examples show that the gestural components are not always stable, but that they reduce to general clusters that vary. This variance depends on which participant is gesturing, as well as on whether the guesser guesses early on during the gestures.

Each set of gestures associated with a component is separated into a single figure. Each figure is labelled with a description of what gesture the participant is making. The numbers of frames are listed per component on the image. After every

\(^{10}\text{This was clear from the interviews that followed the experiment.}\)
set of components per iteration there is a description of the participant’s actions. Each stage of simplification is shown to occur sequentially along the examples.

After sections 3.3.1 and 3.3.2, Chapter 4 expands on the confirmation of the two hypotheses and on answering the research questions as it deals with the simplification of gestures.
3.3.1 - Example: *bag*

PAIR 9 – *BAG*. ITERATION 1. (Participant 9A)

*Figure 26*. Participant 9A mimes opening a bag. The gestures involved create component 1 of iteration 1.
Figure 27. Participant 9A mimes putting objects inside the mimed bag. The gestures involved create component 2 of iteration 1.

Figure 28. Participant 9A mimes closing the bag. The gestures involved create component 3 of iteration 1.
Figure 29. Participant 9A mimes putting the bag (or sack) on his back. The gestures involved create component 4 of iteration 1.
Figure 30. Participant 9A mimes walking with the bag on his back. The gestures involved create component 5 of iteration 1.
Figure 31. Participant 9A mimes putting the bag down. The gestures involved create component 6 of iteration 1.

Figures 26 to 31 show the six components involved in pair 9’s first iteration of bag. All six of the components are iconically representing an agent’s set of actions. Participant 9A pantomimes a scene where he opens a bag, puts objects inside, closes the bag, then puts it over his shoulder and walks around. This is an example of a participant using a longer string of components to convey the concept. These components fit into stage 1:

1 - The gesturer chose to use either one gesture or a series of iconic gestures until the guesser correctly guessed the concept. This was either a single gesture
or a long set of components, depending on how well the gesturer depicted the signified object, how complex the series of actions were, and how well the guesser interpreted the gesture/s.

PAIR 9 – *BAG*. ITERATION 2 (Participant 9A)

*Figure 32.* Participant 9A mimes putting a bag on his back. The gestures involved create component 1 of iteration 2.
Figure 33(1). Participant 9A walks while holding the mimed bag. The gestures involved create the first half of component 3 of iteration 2. The first half of component 3 is seen in Figure 33(2). The gesturer interrupted the action to depict the shape and location of the bag.
Figure 34. Participant 9A depicts the outline and shape of the bag. The gestures involved create component 2 of iteration 2.
Figure 33(2). Participant 9A walks while holding the mimed bag. The gestures involved create the second half component 3 of iteration 2. The first half of component 3 is seen in Figure 33(1).

Figures 32 to 34 show the three components involved in the second iteration of *bag*. The same participant (9A) is gesturing. Two components from the first iteration are maintained, which are the actions of putting a bag over the shoulder and walking with the bag. Participant 9A uses a new component. The gesturer outlines the bag, highlighting that it is the object that he means to convey. This component occurs in the middle of the mimed actions of the participant walking. As the walking action is interrupted by this gesture, the walking action mime is considered to be continuing rather than beginning again. There are fewer components than in the first iteration. Participant 9A chose not to repeat the opening, filling, and closing of the bag. He also did not mime putting the bag down. Even though he omitted four
components, the remaining two components were repeated to keep the gesture similar to those in the first iteration. This matches stage 2:

2 - On the next iteration, the gesturer chose components perceived the most salient and repeated them. At this point, most participants were consciously attempting to keep the gesture as similar as possible to those in the first iteration.

PAIR 9 – BAG. ITERATION 3 (Participant 9B)

Figure 35. Participant 9B mimes putting a bag over her shoulder. The gestures involved create component 1 of iteration 3.
Figure 36. Participant 9B walks with the mimed bag over her shoulder. The gestures involved create component 2 of iteration 3.

Figures 35 and 36 show the two components involved in the third iteration of bag. Participant 9B is gesturing bag for the first time in the experiment. She chose to use the same two components as participant 9A – putting the bag over her shoulder and walking with it. The first component is performed using a different hand and arm
to what the other participant used, which shows that the gesture is not side-specific. Instead, gestures can be performed on either side if it depicts the same action. As this is participant 9B’s first attempt at bag, this is considered as stage 2:

2 - On the next iteration, the gesturer chose components perceived the most salient and repeated them. At this point, most participants were consciously attempting to keep the gesture as similar as possible to those in the first iteration.

PAIR 9 – BAG. ITERATION 4 (Participant 9B)

Figure 37. Participant 9B mimes putting things in a bag. The gestures involved create component 1 of iteration 4.
Figure 38. Participant 9B mimes putting a bag on her shoulder. The gestures involved create component 2 of iteration 4.
**Figures 37 to 39** show the three components involved in the fourth iteration of *bag*. Participant 9B chose to repeat the component of putting objects inside the mimed bag. Due to the nature of the experiment, there were often gaps between iterations of concepts. The participant might have been attempting to make the gestures for *bag* clearer for the guesser. The gestures became clearer as participant 9B selected a component used in iteration 1. While this iteration has more components than the previous one, the third component involves the participant barely taking one step before the guesser calls out a guess. This is seen in stage 3:

3. *As both participants became habituated to the experiment, the gesturer used fewer components with the expectation that the guesser would “get it” after the provided gestures. With both the guesser and the gesturer attempting to achieve the task more quickly, gestures became faster and less precise. The number of components used remained stable or had slight variations (variation came from the participants repeating gestures, participant difference in conceptualization, as well as*
from participants not being able to complete the entire set of components because the guesser responded too quickly).
Figure 40. Participant 9A mimes putting a bag over his shoulder. The gestures involved create component 1 of iteration 5.

Figure 40 shows the only component involved in the fifth iteration of bag. Participant 9A has reduced the components to a single component and the gestures are still understood as referring to bag. The participant did not get a chance to walk with the bag (as the guesser called out a guess), but also had not intended to walk. This is apparent from how he chose to end the component, in particular the locked position of his legs. In previous iterations the component of lifting a bag flowed into walking with it over a shoulder. For the gestures to flow, the participant anticipated the motion and showed signs of moving for component 2 (walking) during component 1 (lifting). In this component, participant 9A shows no signs of preparing to move following the depicting of lifting. This suggests that he was expecting this one component to convey the concept bag. The component in this iteration demonstrates elements of stage 3:

3. As both participants became habituated to the experiment, the gesturer used fewer components with the expectation that the guesser would “get it” after the provided gestures. With both the guesser and the gesturer attempting to achieve the task more quickly, gestures became faster and less precise. The number of components used remained stable or had slight variations (variation came from the participants repeating gestures, participant difference in conceptualization, as well as from participants not being able to complete the entire set of components because the guesser responded too quickly).

11 Note that the participant still looks as if he could start walking, and if the guesser had not guessed so quickly he would have walked. It was clear from his action that he did not intend to walk, unless necessary.
PAIR 9 - BAG. ITERATION 6 (Participant 9B)

Figure 41. Participant 9B mimes putting objects inside a bag. The gestures involved create component 1 of iteration 6.
Constructing Novel Iconic Signs Through Gesture

Figure 42. Participant 9B mimes putting the bag over her shoulder. The gestures involved create component 2 of iteration 6.

Figure 43. Participant 9B walks with the mimed bag on her back. The gestures involved create component 3 of iteration 6.
Figures 41 to 43 show the three components involved in the sixth iteration of bag. Participant 9B has chosen the same three components that she used in the fourth iteration. Though this iteration has more components than the previous iteration, the variation is understandable as each participant is simplifying the gesture individually. It could be that each participant has a conceptually different idea about the gestural scene involved. Each of the three components flowed from one to the next, and the participant prepares their next gestural component while finishing another. This shows that the participant has a well-formed idea of these three components being involved in bag. While there are not fewer components in this iteration, the variation in component count is predicted in stage 3:

3 - As both participants became habituated to the experiment, the gesturer used fewer components with the expectation that the guesser would “get it” after the provided gestures. With both the guesser and the gesturer attempting to achieve the task more quickly, gestures became faster and less precise. The number of components used remained stable or had slight variations (variation came from the participants repeating gestures, participant difference in conceptualization, as well as from participants not being able to complete the entire set of components because the guesser responded too quickly).

In addition, the components used in the sixth iteration are similar in complexity to the participants’ previous attempt in the fourth iteration. This does not show a reduction in spatial or temporal dimensions, except when compared to the components of the first iteration.

Figures 26 to 43 showed all six iterations of bag, gestured by both participants of pair 9. The components, in order of iteration, were 6, 3, 2, 3, 1, and 3. Across the
three iterations that participant 9A gestured, the components were reduced from 6 to 2, and then to 1. Participant 9B used similar numbers of components – 2, 3, and 3 (the additional component was the action of putting objects in the mimed bag). The four stages described in section 3.2 were demonstrated in this example, even though the components did not entirely stabilize.

The next example is of pair 5 participants performing the components for *mountain*. As with this example, after each component a full description will be provided, and the components of each iteration will be used to demonstrate different stages of simplification.
PAIR 5 – *MOUNTAIN*. ITERATION 1 (Participant 5A)

*Figure 44.* Participant 5A depicts the shape of a mountain. The gestures involved create component 1 of iteration 1. Some intermediary gestures between component 1
and 2 have been removed, as the participant took time to think before the next component.

*Figure 45.* Participant 5A repeats the depiction of the mountain’s shape. The gestures involved create component 2 of iteration 1.
Figure 46. Participant 5A depicts something ascending the mountain. The gestures involved create component 3 of iteration 1.

Figures 44 to 46 show the three components involved in the first iteration of mountain. Participant 5A depicts the shape of a mountain using her two parts of her body (both legs and arms are used). The first time, the participant intends it to be a singular component. After some thought, she repeats the depiction of shape but then continues to the third component which is a depiction of something ascending the mountain slope. The guesser correctly guessed the concepts after the third component. The first two components are iconic depictions of shape, while the third is an iconic depiction of an action (something ascending the mountain). The gestures used are predicted in stage 1:

- The gesturer chose to use either one gesture or a series of iconic gestures until the guesser correctly guessed the concept. This was either a single gesture
or a long set of components, depending on how well the gesturer depicted the signified object, how complex the series of actions were, and how well the guesser interpreted the gesture/s.

PAIR 5 - MOUNTAIN. ITERATION 2 (Participant 5B)
Constructing Novel Iconic Signs Through Gesture

D. A. Jones

Figure 47. Participant 5B depicts the shape of the mountain. The gestures involved create component 1 of iteration 2.
Figure 48. Participant 5B depicts something ascending the mountain. The gestures involved create component 2 of iteration 2.

Figures 47 and 48 show the two components involved in the second iteration of mountain. Participant 5B is gesturing the concept for his first time in this experiment. He chose the same two actions that participant 5A had established in the first iteration. This matches what is expected in stage 2:

2. On the next iteration, the gesturer chose components perceived the most salient and repeated them. At this point, most participants were consciously attempting to keep the gesture as similar as possible to those in the first iteration.
PAIR 5 – *MOUNTAIN*. ITERATION 3 (Participant 5B)

*Figure 49.* Participant 5B depicts the shape of the mountain. The gestures involved create component 1 of iteration 3.
Figure 50. Participant 5B depicts something ascending the mountain. The gestures involved create component 2 of iteration 3.

Figures 49 and 50 show the two components involved in the third iteration of mountain. Participant 5B selected the same two components as were previously used. The gestures involved do not simplify spatially or temporally. The components used are still indicative of stage 2:

2 - On the next iteration, the gesturer chose components perceived the most salient and repeated them. At this point, most participants were consciously attempting to keep the gesture as similar as possible to those in the first iteration.
PAIR 5 – *MOUNTAIN*. ITERATION 4 (Participant 5A)

*Figure 51.* Participant 5A depicts the shape of the mountain. The gestures involved create component 1 of iteration 4.
Figure 52. Participant 5A depicts something ascending the mountain. The gestures involved create component 2 of iteration 4.

Figures 51 and 52 show the two components involved in the fourth iteration of **mountain**. Participant 5A selects the same two components that were established in the first iteration. As with the previous iteration, the gestures involved do not simplify spatially or temporally. The components used are still indicative of stage 2:

2. **On the next iteration, the gesturer chose components perceived the most salient and repeated them. At this point, most participants were consciously attempting to keep the gesture as similar as possible to those in the first iteration.**

Please note that the first frame of the first component is included to show that once participants dropped their card into the box, they would sometimes move immediately into the gesture. This blurs the distinction of the gesture’s onset. Both participants move immediately from ‘dropping’ into component one in the next two iterations.
PAIR 5 – MOUNTAIN. ITERATION 5 (Participant 5A)

Figure 53. Participant 5A depicts the shape of the mountain. The gestures involved create component 1 of iteration 5.

Figure 53 shows the only component used in the fifth iteration of mountain. Compared with the first iteration, the participant has preserved the component in the same manner. She has a less wide stance in this iteration, which is a spatial reduction\(^1\). In addition to this, there is only one component used and not two. The depiction of shape is maintained, while the depiction of something ascending the mountain is dropped. This fits with the description of stage 3 given previously:

\(^1\) This spatial reduction, and any spatial or temporal reduction noted in this thesis, are impressionistic only. They are not systematically investigated as a variable in this research, but they would be useful in extensions of this research.
As both participants became habituated to the experiment, the gesturer used fewer components with the expectation that the guesser would “get it” after the provided gestures. With both the guesser and the gesturer attempting to achieve the task more quickly, gestures became faster and less precise. The number of components used remained stable or had slight variations (variation came from the participants repeating gestures, participant difference in conceptualization, as well as from participants not being able to complete the entire set of components because the guesser responded too quickly).

**PAIR 5 – MOUNTAIN. ITERATION 6 (Participant 5B)**

*Figure 54.* Participant 5B depicts the shape of the mountain. The gestures involved create component 1 of iteration 6.
Figure 55. Participant 5B depicts something ascending the mountain. The gestures involved create component 2 of iteration 6.

Figures 54 and 55 show the two components involved in the sixth iteration of mountain. They are the same two components presented in the first iteration. Participant 5B’s posture during the second component is less twisted than in either his previous attempts in iteration 2 and 3. This is a sign of spatial reduction, which fits with the description of stage 3 given previously:

As both participants became habituated to the experiment, the gesturer used fewer components with the expectation that the guesser would “get it” after the provided gestures. With both the guesser and the gesturer attempting to
achieve the task more quickly, gestures became faster and less precise. The number of components used remained stable or had slight variations (variation came from the participants repeating gestures, participant difference in conceptualization, as well as from participants not being able to complete the entire set of components because the guesser responded too quickly).

Figures 44 to 55 showed all six iterations of mountain, gestured by both participants of pair 5. The components, in order of iteration, were 3, 2, 2, 2, 1, and 2. Across the three iterations that participant 5A gestures, the components were reduced from 3 to 2, and then to 1. Participant 5B used the same number of components – 2, 2, and 2. The three stages described in section 3.3 were demonstrated in this example, even though the components did not show many signs of reducing spatially or temporally.

Further research is required to exhaustively investigate the question of how iconic gestures become simplified. The results have shown the simplification of manual gestures does occur, suggested in hypothesis one (section 3.1). Hypothesis two held that participants preferred to represent four different concept types through different iconic parameters (section 3.2). A general process of simplification also occurs, which all participants instinctively followed. Further discussion of the research question is presented in Chapter 4.
4 Discussion

This thesis had two research hypotheses and one central research question. The hypotheses were: 1) that participants will simplify their iconic gestural components following the establishment of mutual understanding; and 2) participants will choose to represent action word types by an agent's action, artefacts by an agent’s action, natural objects by shape, and animates by a referent's typical action. The central research question was “Given the progression of iconicity to symbolicity in graphical research, how do manual gestures become simplified?”. As the majority of ES research investigates the development of graphical signs, this research aimed to incorporate gestural sign development. Because gestural language systems exist, while purely graphical language systems do not, this allows ES to experimentally test closer-to-natural situations of language creation between humans. To form tools of analysis for gesture, research based on Poggi's (2008) work was used. In particular, Poggi’s detailing of iconic parameters for gestures was adopted.

The hypotheses and research question were tested through a gesture experiment, using conventions similar to those in the game of charades. Participants were given 120 concept tokens, which were made up of six iterations of 20 different concepts. Each set of participant pairs was assigned a ‘gesturer’ and ‘guesser’ role. The gesturer would attempt to convey the meaning of a concept through gesture alone, and the guesser would attempt to say the correct concept. After ten concepts had been gestured, participants switched roles. The concept tokens were of four types: actions, artefacts, animates, and natural objects. To analyse the participants’ gestures the notion of gestural components was used. A component is similar to a
syntactic construction in that it is a bundle of smaller parts that create a semantic whole. The number of components per iteration was compared to test hypothesis 1. Hypothesis 2 was tested by coding the participants’ gestures by iconic parameters and comparing them across concept type. Both hypotheses were used to form a speculative answer to the research question, demonstrated through two examples (sections 3.3.1 and 3.3.2).

This chapter is organised as follows: hypothesis 1 is outlined and explored in section 4.1, hypothesis 2 is outlined and explored in section 4.2, and the research question is then discussed in section 4.3. This is followed by the research limitations and research suggestions in section 4.4.

4.1 Hypothesis 1

Hypothesis 1 held that participants will initially construct iconic forms of gesture. Iterations following the establishment of mutual understanding will lead to simplification of the initial iconic form. As discussed in section 4.1, this hypothesis is made up of three different beliefs: 1) participants will gesturally construct the first iteration of a concept with a heavy reliance on iconic forms, 2) when the first iteration is compared with later iterations of the same concept, the level of iconicity will decrease, and 3) iterations following a successful attempt will be simplified. It was expected that all three of these beliefs would be confirmed. The results have shown that participants constructed gestural meaning almost completely by iconic forms, with the exception of a small number of culturally symbolic forms, and deictics. When the components used in the first iteration were compared to those used in later iterations, it was apparent that there was a noticeable decrease in the number of components. This is apparent from the first iteration, with an average
number of 5.9 components (Figure 22, p.60), through to the third iteration, with an average number of 2 components (Figure 22, p. 60). After the third iteration the number of components plateaued, with the fourth, fifth, and sixth iterations all having an average component count of 1.8 (Figure 22, p. 60). When unsuccessful attempts were excluded from the dataset, the number of components still appeared to reduce from the first iteration through to the third. However, the first iteration was frequently spent establishing mutual understanding, and a number of unnecessary gestures were used during this iteration. If the first iteration was excluded, the second iteration still decreased significantly in component number \((t=-2.733, p=<0.001. \text{ Table 7, p. 64})\). Following the third iteration, simplification did not decrease significantly \((t=7.08, p=0.48. \text{ Table 7, p. 64})\). Participants did not reduce their gestured components significantly after mutual understanding was established.

The reason that participants prefer to use iconic gestures when establishing mutual understanding is simple – through iconicity, gesturers are capable of expressing salient features of a signified object, allowing the other interlocutor to conceptualise the object as it is, instead of deciphering information through abstract forms. This is also explained in Poggi’s work: “The very existence of a shared generative device based on iconicity allows for the production and comprehension of newly created gestures. This is why a gesture created from scratch is necessarily iconic: to allow reciprocal understanding between Sender and Addressee” (Poggi, 2008, p. 55). The presence of iconicity in novel gesture is unsurprising, as symbols require conventionalisation. Conventionalisation must necessarily occur on a pre-established sign. It would make sense, particularly for gestural communication, that the pre-established sign began as an iconic sign and then developed into a
conventionalised symbol. Icons never became symbols during this experiment due to the limited number of iterations, which inhibited the process of conventionalisation.

Following the first iteration, the number of gestural components used decreased. This is unsurprising, as during the first iteration participants are in the process of establishing mutual understanding. This could require anywhere between three components to twenty or more components. During this iteration, the gesturing participants must narrow the possibility of which concept they are intending to convey from an entire lexicon down to a specific unit. Frequently, participants even had to distinguish between similar units, for example in the warm-up session participants were asked to convey 'burger'. This was frequently interpreted as 'sandwich', requiring the participant to narrow the meaning even further. The large number of components observed in the first iteration is due to the larger number of gestures required to convey the exact intended meaning. Once mutual understanding was achieved, the additional gestures were no longer necessary, and participants chose to select specific gestural and fewer components to represent the signified object. As hypothesis 2 shows, the chosen components can be partially predicted from the type of concept (action, artefact, animate, or natural object). These chosen components were used each time, leading to minimal reduction once components were settled on. While the second iteration has a higher number of components on average, this is likely due to participants’ partial iteration to help solidify the pairing between the gestural form and meaning. To summarise, participants used a higher number of components to establish meaning during the first iteration. Once the gesture’s meaning was successfully conveyed, participants used a select few of these components to represent the signified object consistently (sections 3.3.1 and 3.3.2).
The number of components that participants used sometimes fluctuated by one or two components during the ‘plateau’ iterations, between iteration 3 to 6. While this was not expected, it was apparent that this was mainly due either to the gesturing participants expecting the guesser to understand the concept when there was a reduction in components or the guesser saying the correct concept before the gesturer had finished fully gesturing. These sorts of fluctuations were never consistent, as participants did not always guess quickly, or leave the implicitly agreed upon set of components ‘unfinished’.

The establishment of mutual understanding is also called ‘grounding’. Garrod et al (2007) investigated the development of graphical signs with conditions that allowed for both more and less interaction during the experiment. They found that communication and interaction were necessary for iconic forms to be simplified. If participants do not interact during the production of the drawing, the drawings “become increasingly complex and retain their iconicity with repeated production” (Garrod et al, 2007: pp 975). The current results are similar in that the gestures became simplified after the successful establishment of mutual understanding. Participants in the graphical experiment simplified their drawings, as shown in Figure 56 which is an example of the graphical simplification during an experiment in which both participants were directors and engaged each other with feedback in the form of drawings (Garrod et al, 2007). The final instance (Block 6) of the concept was represented with a picture that was introduced in the third instance after feedback in the same form (Block 3). In the current research, participants never introduced new components after the successful guess. The difference between simplifications in the two studies is hard to quantify, nevertheless it appears that the graphical
simplifications seem to undergo a greater reduction than the gestural simplifications do.

The final graphical sign is abstracted from an iconic representation of Robert De Niro’s face. The current research’s experiment did not show such a rapid progression of iconicity being simplified. Gestures were simpler at the end of six iterations due to a decrease in necessary components for mutual understanding. The sixth iteration was often an iteration of a single component only, while the first iteration included a variety of novel and repeated gestures.

![Figure 56. The development of the graphical sign for “De Niro”, from Garrod et al (2007).](image)

The pictures in Figure 57 show one participants' gestures for ‘door’ during iteration one. Figure 58 shows the gestures for ‘door’ during iteration six by the same participant.
Figure 57. Participant 6A performs the gestures for ‘door’ (first iteration). This iteration involves 5 components: opening the door (1, 2), outlining the shape (3, 4), opening the door again (5, 6), outlining the shape again (7, 8, 9), and finally opening/closing door (10, 11, 12, 13).
Figure 58. Participant 6A performs the gestures for ‘door’ (sixth iteration). This is the whole component. The participant is miming the act of opening a door.

There is a clear difference in the number of gestures performed in each of the two iterations. Compare the two pictures in figure 58’s with pictures 1, 2, 5, 6, and 10-13 from figure 57. The mimed action is the same, they all represent the act of opening a door. In figure 58, however, the participant uses her right hand instead of her left hand. This allows her to minimize the amount of effort required to perform the action, as she no longer needs to turn her whole body. The action is contained within the movements of the arm and hand. The gestural representations do not change as much as the graphical representations did in Garrod et al (2007). However, a similar development is noticeable. The sixth block in the graphical representation (an arrow which is circled) is a simplified drawing of the four-way directional arrow point in the third block. In the same way that an aspect has been selected and simplified there, the gestures in this experiment were similarly simplified.

The result shows that gestural components will significantly reduce after mutual understanding is established ($t=-11.858, p=<0.001$. Table 7, p. 64). This simplification reduces significantly after the second iteration ($t=-2.733, p=<0.001$. pg. 116
Table 7, p. 64), but plateaus once a functional set of components are selected \((t=7.08, p=0.48, \text{Table 7, p. 64})\). If there were more iterations, a greater deal of simplification might be observable. This would be seen more distinctly in temporal and spatial reductions, instead of gestural component reductions. There is a limit to how much simplification components can undergo. For example you need at least one gestural component to act as a signifier for a signified object. Because of this minimum limit, components can never reduce to be fewer than one component. However, the gesture’s use of space and time should undergo further reduction in addition to the reduction of components. Although components cannot be reduced to fewer than one, simplification is not limited to components. For abstraction, and ultimately the formation of symbolicity, an iconic gesture needs to change. This process must be through simplification as the gesture needs to be used repeatedly and become conventionalised. This type of simplification from iconicity happens because humans prefer to make things more efficient, particularly when it comes to communication. Once both participants understand what message the gestural sign is intended to convey, the sign could then be reduced to allow for easier recognition.

### 4.2 Hypothesis 2

Hypothesis 2 held that participants will primarily choose to represent action word types by an agent’s action, artefacts by an agent’s action, natural objects by shape, and animates by a referent’s typical action. As shown in section 4.2, these patterns, observed originally in Merola (2007), were also observed in the current research. Tables 8 (all iterations) and 9 (iteration 6 only) show that: action words were almost always represented by an iconic agent’s action (all iterations: 90%, iteration 6 only: 91%); natural objects were often represented by an iconic shape (all
iterations: 46%, iteration 6 only: 54%), but were also represented by an iconic referent’s action (all iterations: 23%, iteration 6 only: 25%) and by an iconic agent’s action (all iterations: 21%, iteration 6 only: 15%); animates were very often represented by an iconic referent’s action (all iterations: 66%, iteration 6 only: 70%); and artefacts were most often represented by an iconic agent’s action (all iterations: 60%, iteration 6 only: 66%) but were also represented by an iconic shape (all iterations: 27%, iteration 6 only: 23%). These results match the preferences shown in Merola (2007).

As stated in section 4.1, iconicity is a useful framework for attempting to establish mutual understanding when two interlocutors do not share a communicative system, and further it is a useful framework to solidify meaning; acting as a useful tool when paired with paralinguistic gesture (Birdwhistell, 1952; Merola, 2007). The types of iconicity used to represent different types of concepts could be due to different aspects of salience. These aspects of salience are shown in the results.

Animates are more likely to be represented by their movements. This could be because their most recognisable features are in movement. An alternative explanation is that the animate object’s movement is the easiest aspect to reproduce. Another iconic parameter commonly associated with animates is agent action, for example the gesturer mime s interacting with an animate. This was particularly true for the concepts child and mosquito. For the concept mosquito, participants traced the movements of a mosquito through the air with their index finger (referent’s typical action), and then squashed it when it landed on them (agent’s action). Child was
almost always represented by imitating a child\(^\text{13}\) (referent’s typical action), as well as by shape, with participants indicating the height of a child\(^\text{14}\). *Fish* was represented by a referent action only, where participants either mimed the movement of a fish or the facial motions of a fish. *Caterpillar* was sometimes represented by shape in early iterations, but represented by a referent’s typical actions only during iteration 6. Participants would either crawl on the ground, or ‘crawl’ their index finger along their arm to mean caterpillar. *Dog* was always represented by imitating a dog running around or barking (referent’s typical action). It seems that the movement of an animate object is what gesturers express the most because of the saliency and uniqueness of the movement.

Artefacts are most frequently represented through an iconic agent’s action. This could be because artefacts are most commonly interacted with, so miming the gestures related to the artefact makes sense. If a participant was asked to distinguish between two similarly shaped objects, such as *book* and *chopping board*, the participant would most likely choose to represent the book through miming the action of reading, while representing the chopping board through miming cutting objects on it. Both actions would be accompanied by a deictic, as the participant will most likely want to distinguish the action they are doing and the concept they are attempting to represent. The five artefact type concepts chosen for this experiment were varied in shape, but participants still chose to represent the objects through agent action primarily. Participants often represented *door* by tracing its shape and

\(^{13}\) Participants imitated a child in various ways. Usually, they mimed having a tantrum, but they would also mime an exuberantly happy child.

\(^{14}\) This was usually following another action where participants pointed at themselves. The pair of movements meant “me, but smaller”. Self-referential movements such as this were documented as “Cause-Effect relationship”, which is one of Poggi’s “Medium Referents (MR)”. The total number of gestures marked as cause-effect was very minor and has not been included in the results due to this. These types of MR iconic parameters are expected more in abstract concepts.
then miming the action of opening and closing it, but by the sixth iteration it was always the action of opening the door. Chair was almost always represented by the action of sitting, and then participants would point underneath them. If the guesser failed to guess correctly the first time, the gesturer would either point at the mimed chair again or would depict the shape of the chair. Many participants struggled to convey the concept of boat. The first iteration was frequently the participants depicting the environment a boat would be in, miming the action of driving a boat, or depicting the shape of a boat. The most successful type of gesture was the depiction of a boat’s shape. Nevertheless, participants usually settled on the action of rowing a boat as the representation of boat. Bag was usually depicted by shape first, and then by the action of putting a bag on. If the guessing participant did not successfully guess on the first try, the gesturing participant indicated the bag by pointing at it. Wheel often incorporated all four iconic parameters. Participants depicted the shape, the rotational motion of a wheel, indicated where the wheels were in relation to a car, and they would also mime driving a car. The most logical way to represent artefacts for many participants is through miming the action performed with, or to, the artefact. This could be due to the action being familiar in both the gesturer’s mind and the guesser’s mind; they both have similar experience utilising the given artefact.

Natural objects are most frequently represented by their iconic shape, but this frequency is slightly over half of the total components for natural objects (iteration 6 only: 54%). The next most frequently used type of iconic gesture was a referent’s typical action (iteration 6 only: 25%). Sun was always depicted by shape (i.e., depicting a circle, sometimes with the culturally symbolic “spokes” that are used in children’s drawings of the sun), as well as by agent’s action such as the participant fanning themselves from the heat.
Tree was depicted through either shape or a referent’s typical action only, such as the act of a tree growing from a sprout to a full tree (depiction of growth was classified as both shape and a referent’s typical action) or the movement of the tree’s branches during the wind. During gestures for cloud, participants often referred back to the action for sun. They would depict the sun as too hot, but when the cloud came overhead everything was “okay” again. The most frequently used iconic parameter for cloud was shape, with participants tracing the culturally symbolic cloud shape either on the wall or in the air (the cloud shape commonly depicted was flat on the bottom, and then bubbly/fluffy on the top). Mountain was represented by the iconic conical shape primarily, but participants would also mime hiking up the mountain to solidify the meaning. In some cases, participants would mime the actions of hiking separately to the depiction of the mountain’s shape, and in some other cases the action of hiking was denoted simultaneously to the depicted mountain shape (see section 3.3.2). River was depicted by shape, referent action, and sometimes agent action also. Participants usually attempted to depict the curved shape of a river first, and then they would depict the motion of the water. If this did not successfully convey the meaning, they would interact with the water in various ways. This was anything from being a fish in the river to fishing in the river. By iteration 6, river was usually depicted by a wiggling arm movement only to illustrate the river’s typical motion.

There are many aspects of natural objects, and they vary in many ways. For example, a mountain is immobile, but a river contains constantly moving water. A mountain is far less likely to be depicted by its typical movements because it does not

---

15 “Okay” is used here as the participants frequently used the ‘thumbs up’ gesture to signify a more agreeable state.
have any typical movements, but a river can be depicted by its typical movement. Different types of natural objects will have certain constraints on them, causing gesturers to limit how they depict them. These constraints are all based on what the signified object looks like, how it might move, and our experience related to the object. Shape might be the most commonly used iconic parameter for natural objects due to the most common experience people have with these five natural objects: *mountain, river, tree, sun, and cloud*. The most common experience people have with them is perceiving them through sight, through which we easily determine shape.

Each one of the five natural objects is larger than a human (trees can be smaller but will often grow to be larger). Their size might have an impact on the gestures used to depict them. The objects are too large to easily interact with, so the depiction through an agent’s action is less likely in most cases - particularly with *sun* and *cloud*. Not all the objects move, and their movement that they do have would fail to distinguish the signified object by itself. For example, consider how *river* or *tree* might look through the depiction of their typical action alone. Location was used more during gestures for natural objects than for any other concept type, which may have been to do with the size of the depicted objects. A smaller natural object might cause people to gesture using different parameters. For example, both *grass* and *flower* are smaller natural objects that would be difficult to depict through shape, albeit not impossible. If a gesturer selects parameters based on their personal experience with the object, it would be expected that *grass* could be depicted by location or an agent’s action (mowing the lawn). Similarly, *flower* could be depicted by the gesturer miming a flower in their hand and then smelling it. Another possible reason for the difference is the types of concepts used from the category of natural objects. Merola’s word list for natural objects included *valley, mountain, point,* and
beak. The current research had mountain, river, sun, tree, and cloud. It is possible that within the category of natural objects there are further classifications, for example there are natural objects that are capable of being expressed by either a referent’s or an agent’s actions, and there are those that are not able to be expressed by those parameters.

The final concept is action, which is almost always represented through agent action (iteration 6 only: 91%). To think was depicted by participants miming the act of thinking. To crush was represented by either acting out having a crush on someone, or by actively crushing something. For to find, participants mimed a scene of looking for something, and then actively finding it. To tie was shown by tying an object, whether miming a string being tied, or a shoelace, or a necktie/bowtie. To swing was the only action word that was depicted by something other than agent action. Participants also depicted the shape of a swing set before miming sitting and swinging. By the final iteration, participants only depicted the agent’s action. An easy way to describe actions is to perform them, therefore it is unsurprising that participants chose to represent actions through actions. Participants did not indicate the subject of the verb while gesturing any of the action words, as it was apparent from the mimed context that the gesturer was the subject of the action. This is similar to sign languages in which the omission of a subject is characteristic of a pro-drop (pronoun-drop) language (Padden, 1983; Liddell, 2003; Cormier, Smith, & Zwets, 2010). As the subject of the action was normally contextually clear during
pantomimic gestures, it is not surprising that the subject was never explicitly gestured\textsuperscript{16}.

As mentioned, natural objects might be able to be classified further. This classification would allow predictions of iconic parameters for different types of natural objects. However, this classification does not necessarily need to be restricted to natural objects alone. It was suggested that iconic shape was used based on the size of the natural objects. Because we frequently experience these natural objects visually from a distance, we frequently perceive the outline of the signified object’s shape. For example, a mountain is large, as is the sun, clouds can be big, rivers are always decently long, and trees are generally much bigger than people. As we tend to represent the world in a way that is similar to how we perceive the world, it makes sense that we would represent larger things by shape. The natural objects are far enough away that our most common experience is seeing them from a distance where shape is the most salient feature. The only information we have about the physical world is the information we have gained through our experience and sensory perception. This causes us to represent objects in a manner similar to our perceptual experiences (Bergen, 2012).

Our most common experience with actions is perceiving and performing them. Using a mimed action to represent the action provides a clear and direct iconic link between the signified and the signifier in the mind of the addressee. Occasionally other iconic parameters were used to depict the action, for example \textit{to swing} was also depicted by the participant swinging their arms; representing the motion of swinging

\textsuperscript{16} The only time that the subject (i.e., the participant) was explicitly gestured was during child. Participants often pointed to themselves and then indicated the general height of a child. Instead of the self-referential action being used to explicitly state subject, the action was used to clearly denote the idea of a person – but smaller.
as opposed to the action of swinging (referent’s typical action). The same concept was also occasionally depicted through the iconic shape of a swing set. Normally, iconic shape would not be used to depict an action unless the shape depicts something related to the action.

Artefacts and actions both had 59 agent action component tokens at the sixth iteration (Table 9, p. 68). While agent actions accounted for 91% percent of the action words, they only accounted for 66% of the artefact concepts. Artefacts used more components than any other concept type, with a total of 90 artefact components being used at the sixth iteration. Compare this to 65 action components, 70 animate components, and 79 natural object components. The larger component count for artefacts indicates that participants required more components to represent an artefact. Most of these components were agent action. Participants most commonly depicted an artefact by the actions associated with it, for example door was frequently depicted through the action of opening the door. The participant also used additional iconic types, particularly shape, to indicate that the action (i.e., opening the door) depicted the signified concept (i.e., door). This was an important distinction to make, as a guesser might have guessed that the action concept (i.e., opening/to open) was being mimed, instead of the artefact concept (i.e., door). As our most common experience with artefacts is seeing them or using them, it is again unsurprising that participants commonly represented artefacts by miming the use of them.

Our most common experience with other animate objects is perceiving them and interacting with them. The most frequently used iconic parameter to represent animate objects was referent’s typical action, which is 70% of the total components used to represent animates at the sixth iteration (Table 9, p. 68). This indicates that
the participants preferred to represent an animate object by their typical actions. While a gesturer could use an iconic agent’s action, i.e. miming interaction with the animate object, it could have been difficult to disambiguate between agent action used to depict artefacts or actions. Sometimes participants needed to make the intended signified animate distinct, so they used either iconic shape or an iconic agent action for clarity.

There is a correlation between concept type and iconic parameter selection for representation. This correlation comes down to what type of experience we have had with the signified object, and how frequently we have that experience. If the representation of the signified object presents ambiguity to the guesser then the gesturer must provide additional iconic gestures to clarify the intended meaning. Once gestures were no longer needed for clarification they began to be omitted.

Although the experimental data reflect the results from Merola (2007), there are two unexpected differences. Firstly, participants represented natural objects by iconic shape less frequently than observed in Merola’s experiment. Instead, participants chose to represent natural objects by a referent’s typical action as well as by an agent’s action. Neither of these iconic parameters were observed in use for representing natural objects in Merola’s study. A possible reason for this is that Merola’s dataset had a total number of 89 iconic gestures, compared to 303 gestures in this thesis’ dataset for iteration 6 only (2,542 gestures across all iterations). The difference in the amount of gestural data retrieved could be a factor as to why participants seem to have chosen to use iconic parameters that Merola’s participants did not use.
Another possibility is that the difference in datasets could be due to the difference between iconic paralinguistic gesture and iconic pantomimic gesture. Paralinguistic gesture is co-constructed alongside speech, and its use can be to support what is said or it can provide additional information. On the other hand, pantomimic gesture is used generally in the absence of speech. In this case of this thesis’ experiment, participants were asked to not use any spoken language during their gesturing attempts. Due to this difference between the two types of gesture, it could be that pantomimic gesture requires additional gestures to construct additional information about the signified object. As paralinguistic gesture is used along with speech, it should be expected that the gesture will contain less information in comparison to other gesture types (i.e. pantomimic or linguistic sign) where the meaning is conveyed solely through gesture.

The second unexpected difference between Merola’s study and the current research is the non-iconic types of gestures that were used. The gestures used that were not iconic varied in type; two were cultural pictographs, one was a dance, and the other three were lexical gestures. The two cultural pictographs that were observed involved drawing the sun, particularly with outward lines to represent the light and heat. Two different participants chose to represent the sun in this way. This is a symbol that is culturally embedded enough that it is easily accessible through Microsoft Word’s clipart\textsuperscript{17}, as shown in Figure 59. This cultural embedding is in English, and potentially other cultures also. As it is not shown to be culturally used in all cultures, the term “universal” is not used here.

\textsuperscript{17} The ease of this symbol’s accessibility shows that it is a common enough symbol with a high enough demand. As this is the case, it has been incorporated in the basic shapes offered by Microsoft Word’s clipart. It is one of the few shapes offered that are not purely geometric. The other symbols include an adjustable face and other weather-related symbols.
Two lexical gestures were used to demonstrate aspects of time and sequential ordering. One participant represented caterpillar by first holding an index finger up to represent ‘first’ (Figure 60), and then she depicted the movement of the caterpillar, indicating that the caterpillar is the first stage. Following this, she symbolically represents the future by moving her right hand from her chest further away in front of her (Figure 61), and then depicted the movement of a butterfly. The representation of ‘first’ could have been categorised as iconic, due to one finger indicating a numeric value. However, the gesture was classified as symbolic due to the pre-existing universal use of fingers to represent numbers. The representation of the future through gesture is a temporality marker, and it is a culturally specific representation of time. As temporality is an abstract notion, the gesture was notated as symbolic. The participant who used these symbolic lexical markers also knows ASL, therefore the participant understood how to make gestures lexical.
Figure 60. Participant 7A uses her index finger to indicate the number one, indicating that the following component was the first stage of a caterpillar.

Figure 61. Participant 7A depicts future tense to indicate that the next component is in the future when compared to the previous component, i.e. a butterfly is the future of a caterpillar.

The third lexical gesture was done by a different participant, who gestured a negative lexical gesture. To depict cloud, the participant depicted a cloud, and then rain (Figure 62). To make it clear that rain was not the intended concept, he then traced lines through where he had depicted rain (Figure 62, images 6-8) and made an ‘X’ with his arms (Figure 62, images 10-12). This type of gesture, in conjunction with the previous gestures, could be broadly classified under “relationship of opposition”.
and its negation”, which is a medium referent (MR) based iconic parameter (Poggi, 2008). However, cloud and rain are not opposite to one another. The participant used the negation gesture to inhibit or suppress the guesser from saying rain instead of cloud. As mentioned in Bergen (2012), people hear a negated phrase and first mentally construct a counterfactual simulation before constructing a factual simulation. Participant 8B may have used the gesture of rain to cause the guesser to think of a counterfactual simulation. A benefit of this is that the counterfactual simulation “could be part of the process that allows you to determine what the factual situation actually is in the first place” (Bergen, 2012, p.146). It is possible that participant 8B gestured rain as a contextual clue, as rain and clouds are causally related, but also wanted to clarify that contextual clue was not the intended signified concept. This gestural component by itself is categorised as a lexical marker, rather than as an iconic gesture - especially as negation is not able to be iconically gestured\(^\text{18}\).

\(^\text{18}\) For a gesture to be iconic, it needs to have an aspect of the signified object that can be recreated. Negation is an abstract concept, and the gestural forms associated with negation are culturally derived.
Constructing Novel Iconic Signs Through Gesture

D. A. Jones

Figure 62. Participant 8B depicts rain (1-5), then ‘scribbles’ it out (6-8), and indicates an ‘X’ as a symbol of negation (9-11). Finally, he goes back to gesturing rain to ensure clarity that rain is negated and not cloud.

The final instance of a non-iconic gesture being used was during the concept river, where the participant used dance to represent the concept. The gesturer had tried a few gestures but felt that the meaning was not clear as the guessing participant had made no guesses. He decided to dance, rolling his balled hands around each other (Figure 63, images 1-7), and then depicting the iconic movement of a river while dancing (Figure 63, images 8-15). The dance is representational of lyrics from the song “Proud Mary” by Creedence Clearwater Revival; “rolling, rolling, rolling on a river”. The component gestures match the lyrics on this line, with the rolling fists representing “rolling” and the arm movement as a representation of river. Although both components of the whole dance are iconic in nature, the dance itself is symbolic. The paired participant understood the gesture, and made the correct guess
shortly after seeing the gesture\textsuperscript{19}. Following the establishment of mutual understanding, the participants reduced the sign to one component: an arm curving up and down, iconically mimicking the movement of water (Figure 64). This was the main component of meaning from the dance, and it was the only component part of the first iteration that was repeated.

\textsuperscript{19} Though the paired participant understood quickly, the researcher did not understand the origin of the symbol. After contacting the participant, the dance move was confirmed to be from “Proud Mary”. The gesture was also repeated to a colleague, who understood where it was from immediately. The symbol is clear enough through the dance rhythm and dance gestures that even an intermediary representation of the gesture is understandable to an uninvolved party.
Figure 63. Participant 10B uses a dance from Creedence Clearwater Revival’s “Proud Mary” to represent river during the first iteration.
These six non-iconic types of gesture were not included for quantitative analysis (section 2.4) due to their lack of iconicity. It is apparent that participants are willing to use any type of gesture to communicate. While iconic gestures are prevalently chosen, symbolic gestures are also chosen if the participant feels it can.

*Figure 64.* Participant 10B using only one part of the dance to represent *river* during the sixth iteration.
communicate the correct message. Symbolic signs allow people to easily communicate through a pre-established convention. Participants sensibly utilised symbolic gestures and lexical gestures when the intended message seemed too complex for purely iconic gestures.

Hypothesis 2 has shown the same result as found in Merola (2007); gesturers prefer to use particular types of iconic parameters to represent different concept types. This is captured by Poggi “The aspects selected to be represented manually highly depend on their degree of distinctiveness, on their ease of depiction, and on the ontological type of the target meaning – whether action or entity, animate or inanimate, natural object or artefact”. (Poggi, 2008, p. 55).

4.3 Research Question

The central research question that was investigated was “Given the progression of iconicity to symbolicity in graphical research, how do manual gestures become simplified?” The results from the two research hypotheses show that gestures simplify in terms of components across six iterations of the same concept. This process of simplification is in two essential parts: 1) participants use iconic gestures to represent various aspects of the specific concept, which depended on the participants’ most frequent experiences with the concept. These iconic gestures are usually between five to 10 components, but can be greater (over twenty) or fewer (just a single component); 2) participants implicitly agree on a component gesture, or component gesture set, which is used to convey the concept. This gesture, or gesture set, is kept throughout the remaining iterations. However, this explained process is overly simplistic and misses aspects of simplification that participants employed.
The term ‘simplify’ means to make something easier to do. The ways that things are simplified involve reductions or refinement. For example, reducing the number of components from eight in the first iteration to two in the sixth iteration is a simplification. The number of components required in the first iteration are unnecessary by the final iteration, therefore reduction happens. However, gestures also seemed to be simplified by both temporal and spatial dimensions. Across the course of iterations, while it seemed to stabilise or plateau after iteration 3, participants reduced their gestures by making them faster as well as smaller. While participants appeared to simplify their gestures spatially and temporally, this is not able to be shown without appropriate technology, such as motion sensors which would allow for a clear documentation of both the spatial and temporal dimensions. An example of spatial simplifications is provided in figures 65 and 66. Figure 65 shows the gestures performed for tree in the first iteration, while figure 66 shows the same concept gestured during the sixth iteration. While the movements are similar, the participant does not extend his hands nor arms as far during the sixth iteration. This shows that there are further aspects of simplification that could be analysed in future research, which is discussed more in section 4.4.
Figure 65. Participant 1A performs the gesture for *tree* (first iteration). This is the whole component. The participant is depicting the growth of a tree.

Figure 66. Participant 1A performs the gesture for *tree* (sixth iteration). This is the whole component. The participant is depicting the growth of a tree.

While spatial and temporal simplifications will be easier to document using motion sensor technology, the largest issue is defining the precise boundaries of a component. This type of distinction must be able to identify the onset of a gesture as well as the termination. Spatial and temporal analysis would be difficult without a precise framework that can reliably locate component boundaries. This is discussed further in section 4.4.

Previous work has also shown simplifications occur graphically (Garrod et al, 2007). This research analysed the graphical simplifications by using the Perimetric Complexity (PC) measure. An uninvolved person was asked to judge the graphical complexity of each stage of drawing. It was reported that there was an 82% match between what the person judged more or less complex and what the PC analysis

---

20 This was defined by Attneave and Arnoult (1956), and then developed further by Pelli, Burns, Farrell, and Moore (2006). The formula is PC = Inside + Outside Perimeter² / Ink Area.
showed. This sort of analysis allows for a clear numerical indication of how much ‘ink’ was used in the construction of a drawing. The total surface area of ink could be represented by a volumetric spatial analysis. While this form of analysis would not be immediately useful to gesture, it does demonstrate that gestural ES research does require a form of spatial analysis. Both graphical modality and gestural modality are different in that drawings exist in a two-dimensional space, while gestures exist within a four-dimensional space.

This research has shown several things. Firstly, participants chose to use iconic gestures to depict concepts of specific types. Secondly, the iconic parameters of the gestures could be predicted based on their type. Thirdly, participants reduced the number of gestural components needed for a concept over the course of six iterations. And fourthly, this reduction took place quickly after the establishment of mutual understanding \( t=-11.858, p=0.001 \) (Table 7, p. 64), and continued to reduce after the second iteration \( t=-2.733, p=0.01 \) (Table 7, p. 64), but plateaued after the third iteration \( t=7.08, p=0.48 \) (Table 7, p. 64). Gestural components reduced, or simplified, after mutual understanding was established. This showed one of the ways that gestures simplify across a course of iterations. Spatial and temporal differences were subjectively noticed, and an examination of these two dimensions will allow researchers to answer how an iconic gesture simplifies and abstracts.

---

21 The remaining 18% was examined, and there is a discrepancy between the two – filled areas were perceived by a human observer as more complex than areas that were unfilled.
4.4 Limitations and Future Research

This section lists the limitations within this research, paired with relevant suggestions for future research.

The first limitation was the subjectivity with component boundaries. The component boundaries were decided subjectively by the researcher without a precise measure of when they began and ended. To strengthen the analysis in future, the notion of gestural components would require precisely showing the boundaries of where a component begins and ends. The distinct component boundaries will be advantageous if a four-dimensional analysis of space and time was used. Motion sensor technology would be useful to document this type of gestural research. Both the spatial boundaries and temporal boundaries can be recorded, then compared between iterations and across participants. This would allow for investigation of temporal and spatial reductions.

The second limitation was that the quantity of components was the only factor considered in this analysis of the simplification process of iconic gestures. While the analysis of the number of components showed that simplifications do happen, a more thorough investigation would reveal more about the simplification process. Motion sensor technology would offer quantifiable dimensions of space and time\textsuperscript{22}. For example, full motion sensor technology would allow for spatial analysis of gestures. There are many types of motion sensor technology, and each type has limitations. It is suggested that a non-restrictive type of motion sensor is used, for example Microsoft’s Kinect technology utilises a touchless user interface (TUI). This would be

\textsuperscript{22} These quantifiable dimensions require a precise determination of a component’s onset and termination.
ideal for allowing participants to still naturally gesture without having to hold a motion sensing device. Motion sensor gloves, or even a whole suit, would be advantageous also. This is due to their ability to capture movement in more of a four-dimensional way. As gestures exist within a four-dimensional space, technology that can record within that dimension will be of greater use for analysis of spatial dimensions than a two-dimensional video recording. The temporal simplifications can be recorded as well as the spatial simplifications, allowing for more detailed analytical procedures. In addition to this, the full development of a novel communication system will be recorded with high quality. This is like the ability to record the full development of a communication system in graphical ES research.

The third limitation is the use of pantomimic gesture. The pantomimic gestures used in this research showed little difference in the selection of iconic parameters when compared to the paralinguistic gestures documented in Merola (2007). This indicates that there is a preference of which iconic parameters should be used to represent certain concept types, for example verbs were primarily conveyed through iconic agent action. However, the gestures in Merola’s experiment were used by teachers instructing primary school students. The teacher used gestures along with the spoken word, and the teacher’s selected iconic parameters were analysed. This sort of interaction is one-sided, just like pantomime. Particularly within the setting of a game such as charades, pantomimic communications are performative rather than interactive. While the gesturer was standing, the guesser rarely did more than make guesses for the ten gestured concepts. As shown by the development of Nicaraguan Sign Language, interaction and use are two keys to the development of language. For a gestural ES experiment to experimentally recreate a natural situation of language emergence, such as ISN in Nicaragua, an interactive based activity would
be advantageous. While pantomimic gesture is important to research in relation to the iconicity used and the development of this iconicity, participants would benefit from having open interaction with each other. The benefit would be that gestures will develop and simplify naturally across the course of the interaction, mirroring development of natural sign languages. If participants were given a room of puzzles to solve without using spoken language it would provide them the environment to negotiate meaning freely. The interaction and negotiation could be recorded with motion sensor technology, allowing for participants to move to any point in the experiment space.

The fourth limitation is the small number of iterations. The results of hypothesis 1 showed that after mutual understanding was established, the number of components per iteration essentially stabilised. To further investigate this, another research suggestion is to repeat the experiment with more iterations. The advantage of including more iterations is being able to investigate the extent of the simplification process. As shown in (Garrod et al, 2007), graphical communication abstracts over time. To document the development, motion sensor technology would again be advantageous. It was observed during the current research’s experiment that there was further simplification occurring during the iterations. Analysing the spatial and temporal simplifications as well as the gestural component simplifications will give a more holistic answer to this thesis’ research question: how do manual gestures become simplified?

The fifth limitation was that the experimental cards were potentially visible to the guessing participant, especially if the gesturer accidentally dropped the card or held the card at a bad angle. While this was never an issue during the experiment, it
could be a potential issue for future attempts at the research design. To eliminate the possibility of the guesser seeing the card the room arrangement should be different. Placing the cards behind a screen would mean that the guesser is less likely to see the concept, especially if the gesturer had to walk behind the screen to read the subsequent concept. Another possible solution is using a projector screen to display the concept behind the guesser. These suggestions relate to a charades type experiment, if the experiment was more interaction based the experimental cards would not have to be necessary.

More gestural research is needed in the field of ES. Gestural ES research is required to postulate theories of the original development of gesture and communication. The development of gestural communication systems heavily involves interaction and negotiation, both of which are important in all language development. Interaction and negotiation are important to fully document the development from iconicity to symbolicity in human communication.
5 Conclusion

This research aimed to confirm two different hypotheses as well as a central research question. Hypothesis 1 stated that participants will simplify their iconic gestural components following the establishment of mutual understanding. Hypothesis 2 stated that participants will choose to represent action word types by an agent’s action, artefacts by an agent’s action, natural objects by shape, and animates by a referent’s typical action. The main research question for this thesis was “how do iconic manual gestures become simplified?”.

The experiment for this study was similar to charades. Participants were organised into pairs and given 120 concept tokens, comprised of 6 iterations of 20 concepts. Participants were allocated to a ‘gesturer’ or ‘guesser’ role, which switched after ten concepts. The gesturer was asked to convey each concept through gesture alone, and the guesser was asked to guess what the gesturer was conveying. Throughout this process, the participants were video recorded for analysis.

Hypothesis 1 results showed that there is a reduction in the number of components used depending on how many times the concept had been repeated. The largest amount of simplification occurred after iteration 1 ($t=11.858, p=<0.001$), with some simplification occurring after iteration 2 ($t=2.733, p=<0.001$), but not after iteration 3 ($t=7.08, p=0.48$. See Table 6 for complete results). Hypothesis 2 showed that the findings in Merola (2007) and Poggi (2008) are repeatable. The results matched the previous research’s findings, showing that people have preferences for which iconic parameter they use depending on the concept’s type. The research question found that gestures simplify in more ways than by the number
of components used. It was subjectively noticed that the gestures simplified temporally and spatially also.

This research has provided a new research paradigm to the field of Experimental Semiotics. As the research for ES has all been conducted on iconicity within the graphical modality, this research is novel. The results of this experiment have shown that the use of iconicity in gesture can be experimentally studied and documented. In addition, the experiment has shown the usefulness of pantomime for the investigation of iconicity within gesture. Investigating the simplification of iconic gestures allows researchers to understand how signs become symbolic, and stronger theories surrounding the origin of language can be developed. This is especially important to strengthen the well-reasoned hypothesis of language's gestural origin.

Five limitations were stated in section 4.4. The first limitation mentioned was the subjectivity of component boundaries. In order to measure temporal and spatial reductions in gestures, a clear onset and termination needs to be identified. The second was that only component number was investigated. The simplification process would benefit from spatial and temporal analysis in addition to the analysis of component reduction. The third limitation was that the experiment used pantomimic gesture. Though pantomimic gesture is perfect for investigating iconicity, simplifications would occur faster and more naturally if the experiment involved interactive gesture instead of performative gesture. The fourth limitation is the number of iterations. The number of iterations was sufficient for answering this study's research question; however further reductions in component count might be observable across a larger number of iterations. The final limitation is that the guessers were potentially able to see the gesturers' cards. While this was never an
impact on the experiment, any future attempt at recreating this research should consider putting the cards behind a screen so that the guesser cannot accidentally see them.

As this research is novel for the field of ES, there is much more that could be done to expand on what was found. Three main research suggestions were provided. The first suggestion was that motion sensor technology could be used to capture the spatial parameters that gestures use. This would allow for documentation and analysis of simplifications that occur in the spatial dimension. For this type of research to be successful, the precise boundaries of components (i.e. their onset and termination) would need to be clearly distinguished and marked. The second research suggestion is to use an interactive experiment instead of pantomime. While pantomime was useful to investigate iconic gestures, future research could use an interactive experiment that asks participants to use established gestures to accomplish some task beyond recognition. In such a task, simplification would be potentially accelerated in a more natural mode of communication. For example communication that involves interaction as seen in the development of Nicaraguan Sign Language. The third research suggestion was to repeat the experiment in this research but with more iterations. This would reveal how much gestural components can become simplified.
References


pg. 152


