


Relative difficulty of early grade compare type word problems: Learning from the case of isiXhosa

**Author:**Ingrid E. Mostert¹ **Affiliation:**

¹Centre for Education Practice Research, Faculty of Education, University of Johannesburg, Soweto, South Africa

Corresponding author:

Ingrid Mostert,
ingridmostert@gmail.com

Dates:

Received: 03 Mar. 2020

Accepted: 26 Aug. 2020

Published: 21 Dec. 2020

How to cite this article:

Mostert, I.E. (2020). Relative difficulty of early grade compare type word problems: Learning from the case of isiXhosa. *Pythagoras*, 41(1), a538. <https://doi.org/10.4102/pythagoras.v41i1.538>

Copyright:

© 2020. The Authors.
Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Word problems form an important part of the early grade mathematics curriculum in South Africa. Studies have shown that the relative difficulty of word problems differ: learners are more likely to solve certain types of word problems than others, with compare type problems being the most difficult. In order to help early grade learners understand and solve compare problems, it is important to understand the relative difficulty of different types of compare type problems and the factors that contribute to their relative difficulty. While these factors have been studied in English, less research has attended to word problems in other languages, such as isiXhosa. In this study a typology of isiXhosa compare type (difference unknown) word problem was set up. The typology included two dimensions, namely the problem situation and the comparative question. The relative difficulties of specific word problems from this typology were compared by analysing the results from an early grade mathematics assessment administered to two cohorts of Grade 1–3 isiXhosa learners in five rural Eastern Cape schools. The analysis showed that in isiXhosa, as in English, some compare type problems are easier to solve than others. Problems with ‘matching’ situations are easier to solve than problems with ‘no matching’ situations. Problems with alternatively formulated comparative questions, specifically those using *-shota* or *kangakanani*, are easier to solve than those using a more classic formulation. This study highlights the importance of understanding the ways in which African languages express mathematical ideas in order to identify and leverage affordances for teaching and learning mathematics.

Keywords: word problems; early grade mathematics; compare problems; isiXhosa; mathematical language.

Introduction

Word problems are a central, yet hard-to-teach, aspect of early grade mathematics. For example, in South Africa word problems have been identified as a recurring weakness in the South African Annual National Assessments (ANAs) (Department of Basic Education, 2012, 2014, 2015). Research has shown that the relative difficulty of word problems differs: learners are more likely to solve certain types of word problems than others. For additive relation word problems, in other words any word problems involving addition and subtraction, compare type problems have been shown to be the most difficult for learners to solve. Compare type problems are of the form ‘Sbu has eight bananas and Sive has five bananas. How many more bananas does Sbu have than Sive?’ While there has been some research into early grade word problems in South Africa (e.g. Petersen, McAuliffe, & Vermeulen, 2017), and some research into word problems and African languages in higher grades (e.g. Sepeng, 2013), there has been little research into early grade word problems in African languages. This is problematic as more than 75% of learners are taught mathematics in an indigenous African language in the first four years of formal schooling (Spaull, 2016).

Internationally, different types of additive relation word problems and their relative difficulty (as measured by the percentage of learners who correctly answered the problem out of the total number of learners who were asked the question) have been studied in relation to English since the late 1970s. This work was pioneered by two research groups: the first led by Carpenter, Hiebert and Moser (1981) and the second by Riley, Greeno and Heller (1983). Such studies have shown that there are a number of factors that influence the relative difficulty of different word problems. These include general factors such as problem length, grammatical complexity and whether learners use concrete aids or not, as well as specific factors such as semantic structure and the position of the unknown (Riley et al., 1983). Many of these factors relate to language. This raises questions regarding the extent to which these factors influence the relative difficulty of word problems in languages other than English, especially languages with linguistic features

Read online:

Scan this QR code with your smart phone or mobile device to read online.

significantly different to English, such isiXhosa, one of South Africa's nine official African languages.

This study contributes to the limited research on early grade word problems in languages other than English by examining compare type word problems in isiXhosa. The study offers a sub-typology for compare type (difference unknown) problems, both in English and in isiXhosa. The relative difficulty of the problems in the isiXhosa sub-typology is then empirically tested using data collected from an adapted early grade mathematics assessment (EGMA).

The following two research questions are answered in this study:

- Are certain types of isiXhosa compare type problem easier for learners to solve than others?
- If so, is the relative difficulty influenced by:
 - the formulation of the comparative question
 - the problem situation?

Theoretical and methodological perspectives

This article is informed by theoretical perspectives and methodological tools from linguistics that have proved helpful for research into the way different languages express mathematical concepts, as proposed in a recent paper by Edmonds-Wathen (2019). This article also draws on theoretical perspectives from variation theory, a general theory of learning largely developed by Marton and Booth (1997) and later extended by Watson and Mason (2005) in relation to mathematics learning.

Edmonds-Wathen (2019) proposes using a typological framing for research on the diversity of mathematical expression in different languages and using interlinear morphemic glossing to present examples in different languages. These perspectives and methodologies are particularly pertinent for studies done by a researcher not fluent in the language that is being studied. Edmonds-Wathen points out that linguists often work with languages that they are unfamiliar with, either by working with translated texts or by working closely with bilingual speakers. She argues that mathematics education researchers can, and do, work in similar ways, with this study being a case in point. This study was undertaken by an English speaker with an emergent understanding of isiXhosa. The researcher worked very closely with a number of isiXhosa speakers to deepen her understanding of isiXhosa, particularly in relation to compare type problems.

Typological framing

Typology is an area of linguistics that describes and classifies languages according to their structural similarities and differences (Edmonds-Wathen, 2019). Typology strives to compare languages through an analysis that is framework-neutral (Nichols, 2007). Edmonds-Wathen (2019) argues

that because of this neutrality, 'a typological approach may be useful to investigate mathematical expression in different languages, without privileging one language over another' (p. 121). This is particularly important when comparing a language with a well-developed mathematical register (see Halliday, 1978), for example English, with a language without a formal mathematical register or with a mathematical register that is still being strengthened, for example isiXhosa.

While this study does not adopt a strict typological approach, the study does strive to ensure that English was not privileged over isiXhosa, for example by ensuring that a range of ways of expressing comparative questions in isiXhosa was studied and not only those that correspond to the way in which English expresses comparative questions. However, as the researcher is not an isiXhosa speaker, English and the linguistic features used to express comparative questions in English provided a starting point for the study, therefore implicitly privileging English.

Interlinear morphemic glossing

One of the challenges of researching how different languages express mathematical ideas is how to present examples from a language different to the language of publication. To overcome this challenge, Edmonds-Wathen (2019) suggests using a simplified interlinear morphemic gloss.¹ Interlinear morphemic glossing allows the structure of the example to be presented. Often the structure is lost if only an idiomatic translation is provided (Edmonds-Wathen, 2019). Interlinear morphemic glossing is particularly helpful when presenting data from languages where word order is not necessarily a determinant of the function of a word.

An interlinear morphemic gloss consists of four levels (Edmonds-Wathen, 2019). For this article the top level gives the isiXhosa in sentence form. The second level gives the isiXhosa morphemes (the smallest unit of a language that has its own meaning), the third level gives the English morphemic gloss, and the final level gives a free translation in English. Morphemes are separated by a hyphen. If a morpheme is translated by more than one word the words are separated by a full stop. For example:

Level 1: Umama uneembiza ezisibhozo neziciko ezihlanu.

Level 2: *umama u-nee-imbiza ezi-sibhozo ne-ziciko ezi-hlanu*

Level 3: mother she.is-with-pots that.are-eight with-lids that.are-five

Level 4: Mother has eight pots and five lids.

The simplified interlinear morphemic gloss is used in order to make the examples accessible to a mathematics education audience. For this reason, in some instances in this article, not every morpheme is glossed separately.² For example, nouns and their prefixes, which indicate the noun class and the

1. See Comrie, Haspelmath and Bickel (2008) for the Leipzig glossing rules used widely in linguistics.

2. IsiXhosa is an agglutinative language and words are made up of many morphemes.

number (singular or plural) of the noun, are not glossed separately. In cases where strict glossing would detract from the comparisons being made, other glossing rules have also not been strictly followed.

Language of variation

Variation theory is a general theory of learning that has been applied specifically to learning mathematics. A central notion within variation theory is that in order to discern one aspect of a phenomenon, that aspect needs to be varied while other aspects remain unchanged (Al-Murani, Kilhamn, Morgan, & Watson, 2019). The aspect of the phenomenon that varies is called the 'dimension of variation'.

Watson and Mason (2005) extended these ideas by defining the variation that is possible within a 'dimension of variation' as the 'range of change'. For example, in the expression $x + 3$, one of the dimensions of variation is the addend (others include: the letter representing the variable, the operator and the order in which the variable and constant appear in the expression). The values that the addend can take (i.e. natural numbers, negative numbers, rational numbers and so on) constitute the range of change of this particular dimension of variation (Al-Murani, 2006). The extent to which a learner can discern the dimensions of variation and the corresponding 'range of change' of the expression $x + 3$ is an indication of how well the learner understands the algebraic expressions.

In this article the ideas of a 'dimension of variation' and of a 'range of change' are applied not to an object of learning, but to an object of study, namely compare type word problems. In order to explore the full range of possible compare type problems, different dimensions of variation were identified and varied to set up a typology of compare type problems for isiXhosa and for English.

Additive relation word problems

Additive relation word problems and the factors that influence their relative difficulty have been studied for English word problems since the late 1970s. In the following two sections relevant studies are discussed.

Word problem typologies

Early researchers categorised word problems describing the same mathematical problem but using different semantic structures into typologies of word problems (e.g. Carpenter & Moser, 1983; Riley et al., 1983). Recently these typologies have been combined into one comprehensive typology (Mostert, 2019). The categories and labels from this comprehensive typology will be used in this article. At the highest level this typology consists of four different types of word problems, differing in terms of the number of sets being compared and whether the problem is dynamic or static (Figure 1).

Each of these four main categories of word problems can be separated into subcategories in two ways. Firstly, each category can be separated into two subcategories based on a number of factors or dimensions: the 'direction' of the change or equalisation, whether attributes or ownership are different in collection problems and whether the comparison is 'more than' or 'less than'. Secondly, by changing the position of the unknown, each category can further be divided into two or three subcategories resulting in a total of 22 subcategories (see Figure 2).

The word problems studied in this article are 'more than' compare problems where the difference is unknown (marked with † in Figure 2).

Factors influencing difficulty of word problems

As mentioned previously, there are a number of different factors that have been identified as influencing the difficulty level of word problems. The factor that is relevant for this study is the clarity of the problem for learners, both in terms of the problem situation and in terms of the formulation of the comparative question. Verschaffel and De Corte (1993) report that most mistakes made by young children when solving additive relation word problems are more likely to be because they represent the problem situation incorrectly, not, as was formerly widely believed, because they choose the incorrect arithmetic operation. This is evident from a number of studies, referred to and validated by Verschaffel and De Corte (1993), which demonstrate that problems can be rephrased, without

	Dynamic (like a movie)	Static (like a photo)
Single set	Change A girl has five sweets. Then she gets two more sweets. How many sweets does she have now?	Collection A girl has five red sweets and two blue sweets. How many sweets does she have altogether?
	Equalise A girl has seven sweets. A boy has five sweets. How many more sweets does the boy need to have the same number of sweets as the girl?	Compare A girl has seven sweets. A boy has five sweets. How many more sweets does the girl have than the boy?
Separate sets		

Source: Adapted from Mostert, I. (2019). Distribution of additive relation word problems in South African early grade Mathematics workbooks. *South African Journal of Childhood Education*, 9(1), 1–12. <https://doi.org/10.4102/sajce.v9i1.655>

FIGURE 1: Four main categories of word problems.

Change		Collection	
<i>Increase</i>	<i>Decrease</i>	<i>Attributes</i>	<i>Ownership</i>
Result unknown	Result unknown	Collection unknown	Collection unknown
Change unknown	Change unknown	Subset unknown	Subset unknown
Start unknown	Start unknown		
Equalise		Compare	
<i>Increase</i>	<i>Decrease</i>	<i>More than</i>	<i>Less than</i>
Change unknown	Change unknown	Difference unknown†	Difference unknown
Target unknown	Target unknown	Compared quantity unknown	Compared quantity unknown
Start unknown	Start unknown	Referent unknown	Referent unknown

Source: Summarised from Mostert, I. (2019). Distribution of additive relation word problems in South African early grade Mathematics workbooks. *South African Journal of Childhood Education*, 9(1), 1–12. <https://doi.org/10.4102/sajce.v9i1.655>

†, The word problems studied in this article are ‘more than’ compare problems where the difference is unknown.

FIGURE 2: Twenty-two subcategories in the comprehensive typology.

changing their semantic structure, in way that is easier for learners to correctly represent the problem situation.

For collection problems there are two cases of rephrasing that have been shown to increase the likelihood of learners solving the problem correctly. Firstly, Carpenter et al. (1981) found that if the problem ‘There are six children on the playground. Four are boys. How many are girls?’, was changed to ‘There are six children on the playground. Four are boys and *the rest are girls*. How many are girls?’, a higher percentage of learners answered the question correctly.

Secondly, Lindvall and Ibarra (1980) found that when the problem ‘Together Tom and Joe have eight apples. Three apples belong to Tom. How many belong to Joe?’ was changed to ‘Together Tom and Joe have eight apples. Three *of these* apples belong to Tom. How many *of these* belong to Joe?’, the problem was significantly easier for kindergarten children to solve correctly.

Such rephrasing is also possible for compare type problems, with some empirical data showing that rephrasing can increase the percentage of learners who correctly solve the problem. This will be discussed in detail in the next section.

Compare type problems in English

Compare type problems have been identified as the most difficult type of additive relation word problem for young children to solve (Fuson, Carroll, & Landis, 1996). At least part of the reason why learners struggle to solve compare type problems is because standard compare type problems (‘Sbu has eight bananas and Sive has five bananas. How many more bananas does Sbu have than Sive?’) include quantifiers such as ‘more than’ and ‘less than’. Quantifiers form part of later-developing languages skills, skills that mother tongue speakers continue to develop up to approximately age 9 (Berman, 2004). If children are still learning to understand and use words such as ‘more than’

and ‘less than’ it is not surprising that they struggle to represent the correct problem situation for compare type problems.

Another reason why compare problems are so difficult for young learners to solve is that learners confuse the ‘classic’ comparative question ‘How many more?’ with the question ‘How many?’ or with the question ‘Who has more?’. For example, the problem ‘Sbu has eight bananas and Sive has five bananas. How many more bananas does Sbu have than Sive?’, is often answered with ‘eight’ (correctly answering the question ‘How many bananas does Sbu have?’) or with ‘Sbu’ (correctly answering the question ‘Who has more bananas?’) (Roberts, 2016). In response to this potential confusion, Roberts (2016) suggests first asking ‘Who has more bananas?’, then ‘How many bananas does Sbu have?’, before asking ‘How many more bananas does Sbu have than Sive?’.

In terms of comparison, both mathematics education and linguistic research (e.g. Kennedy, 2009) has focused on ‘more than’ compare type problems, neglecting ‘less than’ compare type problems (see Figure 2). This focus on ‘more than’ problems in the literature is reflected in assessments, such as the EGMA, which only includes ‘more than’ compare type problems. Because this study analyses data from the EGMA, only ‘more than’ compare type problems are considered. This is a limitation of the study and of research in general as ‘more than’ compare type problems are not necessarily equivalent to ‘less than’ compare type problems, especially for non-Indo-European languages. IsiXhosa is one such language where they are not equivalent.

It is also important to note that while there are three subcategories of compare type (more than) word problems, as can be seen from Figure 2 and as exemplified in Table 1, this study only considers ‘difference unknown’ compare type problems.

TABLE 1: Exemplification of subcategories of compare (more than) problems.

Subcategory	Example
Difference unknown	A boy has nine sweets. A girl has eleven sweets. How many more sweets does the girl have than the boy? [†]
Compared quantity unknown	A boy has nine sweets. A girl has two more sweets than the boy. How many sweets does the girl have?
Referent unknown	A boy has some sweets. A girl has two more sweets than the boy. The girl has eleven sweets. How many sweets does the boy have?

Source: Extracted from Mostert, I. (2019). Distribution of additive relation word problems in South African early grade Mathematics workbooks. *South African Journal of Childhood Education*, 9(1), 1–12. <https://doi.org/10.4102/sajce.v9i1.655>

[†], The word problems studied in this article are ‘more than’ compare problems where the difference is unknown.

Different types of compare type problems

In this section, three variations of the standard, ‘difference unknown’ compare type problem are discussed.

In as early as 1980, Hudson constructed and tested a variation of the standard compare type problem. Hudson’s (1980) variation differed from the standard compare type problem in two ways. Firstly, the problem situation was set up to invoke the idea of matching by choosing birds and worms as the subjects of the story, namely ‘There are five birds and four worms’. Secondly, the ‘classic’ phrasing of the comparative question ‘How many more birds than worms are there?’, was rephrased as ‘Suppose the birds all race over and each one tries to get a worm! Will every bird get a worm? How many birds won’t get a worm?’.

Hudson (1980) then presented learners with the same problem situation about the five birds and four worms, but posed the comparative question in two different ways, using the classic ‘how many more’ formulation and using the ‘how many won’t get’ formulation. Table 2 shows the striking difference in results with a much higher percentage of learners being able to solve the problem with the ‘how many won’t get’ formulation.

A second variation was introduced by Roberts (2016), in what she refers to as ‘compare (matching) problems’. These are compare type problems that draw attention to the absence of elements by asking ‘How many elements are missing?’. An example of Roberts’s compare (matching) problem is: ‘There are 11 locks but only 9 keys. How many keys are missing?’. Roberts explains that:

the choice of locks and keys is deliberate, as in this problem context it is implicit that each key fits uniquely with a particular lock. This unique 1:1 matching of each element in one set to each element in another set is not explicitly implied in the compare problem ‘I have 11. You have 9. How many more do you have than me?’ (p. 68)

The unique one-to-one matching of Roberts’s compare (matching) problem is *embedded* in the problem situation. As Roberts (2016) points out, this is not the case for ‘standard’ compare type problems. It is also not the case for Hudson’s ‘won’t get’ problems, where, unlike with locks and keys, one bird can get two worms or two birds share one worm. In Hudson’s (1980) variation, the one-to-one matching is *imposed*

TABLE 2: Percentage of children with consistent correct responses.

Grade	How many more? (%)	How many won’t get? (%)
Nursery school	17	83
Kindergarten	25	96
First grade	64	100

Source: Reproduced from Riley, M., Greeno, J., & Heller, J. (1983). Development of children’s problem-solving ability in arithmetic. In H.P. Ginsburg (Ed.), *The development of Mathematical thinking* (pp. 153–196). New York, NY: Academic Press.

TABLE 3: Percentage of children with consistent correct responses.

Assessment	How many more? [†] (%)	How many missing? [‡] (%)
Cycle 2 pretest	52	68
Cycle 3 pretest	34	73

Source: Collated from Roberts, N. (2016). *Telling and illustrating additive relations stories: A classroom-based design experiment on young children’s use of narrative in mathematics*. Johannesburg: University of Witwatersrand. <https://doi.org/10.13140/RG.2.1.1092.6964>

[†], Referred to as Compare (disjoint set) by Roberts (2016).

[‡], Referred to as Compare (matching) by Roberts (2016).

Type	Example
Standard	A girl has seven sweets. A boy has five sweets. <u>How many more</u> sweets does the girl have than the boy?
Roberts (2016)	A mother has seven pots and five lids. <u>How many</u> lids are <u>missing</u> ?
Early grade mathematics assessment (EGMA)	A mother has seven children and five oranges. <u>How many</u> oranges are <u>still needed</u> if the mother wants to give each child one orange?
Hudson (1980)	There are five birds and four worms. <u>Suppose the birds all race over and each one tries to get a worm!</u> <u>How many</u> birds <u>won’t get</u> a worm?

Note: Question words are underlined, and phrases that impose one-to-one matching are highlighted.

FIGURE 3: Different types of compare type problems in the literature.

on the problem situation by adding the phrase ‘each [bird] tries to get a worm’.

While Roberts (2016) did not use the same problem situation, she did empirically determine the facility score of a compare (matching) problem and a standard compare type problem (which she refers to as a compare (disjoint set) problem). For both pretests in her study the facility score of the compare (matching) problem was much higher than that of the standard compare type problem (see Table 3).

A third variation of the standard compare type problem appears in the EGMA used in this study (Figure 3). This variation is similar to Hudson’s ‘won’t get’ variation. Firstly, the problem situation is set up to invoke matching, this time between children and oranges: ‘A mother has seven children, and she has two oranges’. Secondly, the need for one-to-one matching is *imposed* on the situation rather than embedded in the situation. This is done through the phrase ‘if the mother wants to give each child one orange’. However the EGMA variation differs from the Hudson (1980) variation in that rather than asking a ‘won’t get’ question, a ‘still needed’ question is asked: ‘How many oranges are still needed?’.

While the three variations of the standard compare problem, summarised in Figure 3, are helpful in showing that rephrasing a question can influence how easy or difficult it is for learners to solve, the problems differ both in terms of the problem situation and in terms of the formulation of the comparative question. This means that it is not possible to isolate the effect of the different factors on the level of difficulty of the different problems. In the next section a typology of compare type problems is set up, taking into account the variation that is possible for both factors.

Typology of English compare type problems

Drawing on the language of variation theory, the 'dimensions of variation' for compare type problems are (1) the problem situation and (2) the formulation of comparative question. The problem situation can either be one that invokes matching by referring to two things that learners might expect to go together (e.g. locks and keys or children and oranges) or one that does not invoke matching by referring to things that do not necessarily go together (e.g. sweets belonging to a girl and sweets belonging to a boy). For matching problems, it possible to further differentiate between problems that have one-to-one matching embedded in the situation and those in which the one-to-one matching is not embedded.

In English, the comparative question, which constitutes the second dimension, can either be formulated in the 'classic' form, 'how many more?' or, for matching situations, the question can be formulated in one of a number of alternative ways such as 'how many are missing?' or 'how many are still needed?'

Using these two dimensions of variation and the range of change that is permissible for each dimension, it is possible to set up a typology of compare type problems. Figure 4 provides an overview of the typology as well as showing how each of the four compare type problems discussed previously (see Figure 3) fits into the typology. Appendix 1, Figure 1-A1 exemplifies each of the categories in the typology.

There are few important things to note about the typology.

Firstly, while it is possible to ask a classically formulated comparative question with a matching problem situation (e.g. 'A mother has eight pots and five lids. How many more pots are there than lids?'), it is not possible to use an alternatively phrased question with a 'no matching' problem situation – the problem 'A girl has seven sweets. A boy has five sweets. How many sweets are missing?' does not make sense. This is the reason for the *n/a* cell.

Secondly, when a problem has a matching problem situation where one-to-one matching is not embedded but an alternatively phrased comparative question is used, an additional phrase (such as 'each bird tries to get a worm') must be added in order to impose the one-to-one matching on the situation. For this reason, the typology differentiates

between '1-to-1 matching not embedded' and '1-to-1 matching imposed'. See Appendix 1, Figure 1-A1 for examples of word problems in the different categories. Finally, it is important to remember that this typology is only for 'difference unknown' compare type problems.

Once a broader typology has been set up showing the dimensions along which the problems can vary, it is possible to compare problems that only vary in terms of one dimension (either the problem situation or the comparative question) in order to establish the extent to which each factor influences the relative difficulty of compare type problems. In this study the influence of these two dimensions is explored for isiXhosa compare type problems. In order to do this a typology for isiXhosa compare type problems is set up, drawing on examples from canonical texts.

Compare type problems in isiXhosa

In a previous study, Mostert and Roberts (2020) describe the linguistic features of comparative phrases in isiXhosa. In order to do this they analysed the examples of comparative phrases appearing in four canonical texts,³ written in English and translated into isiXhosa (Mostert & Roberts, 2020). This set of examples included both comparison phrases (e.g. 'There are more dogs than cats') as well as comparative questions (e.g. 'How many more dogs are there than cats?'). While the previous study only focused on the comparison phrases, this study focused on the comparative questions in these canonical texts, while also including comparative questions from the EGMA.

As in the previous study, the isiXhosa texts provide a valuable source of examples of how to formulate comparative questions in isiXhosa, but are not a sufficient source of examples. Because of the small number of examples and because, as mentioned previously, the author is not fluent in isiXhosa, mother tongue isiXhosa speakers were consulted to clarify and exemplify the range of possible formulations of comparative questions in isiXhosa. Before setting up a typology of isiXhosa compare type problems, aspects of isiXhosa grammar that are relevant for the study are discussed.

Relevant isiXhosa grammar

IsiXhosa is a Nguni language spoken by more than 8 million South Africans (of a total of 57 million). The other three Nguni languages spoken in South Africa are isiZulu, isiNdebele and siSwati. As a Bantu language, isiXhosa has many linguistic features that differ substantially from the linguistic features of Indo-European languages. Two features that are relevant for this study are flexibility of word order and a system of concordial agreement.

IsiXhosa word order is not as rigid as English word order. In isiXhosa the most important word in a sentence is emphasised

³The Curriculum and Assessment Policy Statement (Department of Basic Education 2011) and three sets of learner workbooks produced by the national Department of Basic Education (2018), the National Education Collaboration Trust and the Nelson Mandela Institute.

Comparative question	Classic	Problem situation		
		No matching	Matching	
			1-to-1 embedded	1-to-1 not embedded
How many more?	<i>Standard</i>	-	-	
Alternative	No matching	1-to-1 embedded	1-to-1 imposed	
		-	-	
		<i>Roberts (2016)</i>	-	
How many missing?	-	-	-	
How many still needed?	<i>n/a</i>	-	<i>EGMA</i>	
How many won't get?	-	-	<i>Hudson (1980)</i>	

EGMA, early grade mathematics assessment; n/a, not applicable.

FIGURE 4: Typology of English compare problems (including examples from literature).

by putting it at the beginning of the sentence. Table 4 shows how the English sentence 'Sigqibo gave Mveli bread' can be constructed in four different ways in isiXhosa, each emphasising different words. This flexibility of word order accounts for some of the variation of isiXhosa compare type problems.

isiXhosa, like other Bantu languages, has a noun class system. This means that all nouns belong to a particular class which is determined by the noun's prefix. In a sentence, any word (verb, noun, pronoun or adjective) associated with a noun has to show 'agreement' with that noun. This is achieved by adding a concord (a prefix) to the word, which contains similar-sounding letters to the prefix of the noun. This is referred to as concordial agreement. For example:

Izinja zininzi kuneekati. 'There are more dogs than cats.'
Abantwana baninzi kunoomama. 'There are more children than mothers.'

In the first example, the noun is *izinja* 'dogs' with the prefix *izi-* while in the second sentence the noun is *abantwana* 'children' with the prefix *aba-*. In each sentence the adjective *-ninzi* 'many' takes a different prefix, as determined by the noun it is describing. For this reason, when referring to words on their own (i.e. when they are not referred to as part of a sentence), the root of the word is used (e.g. *-ninzi*) rather than one particular form of the word (e.g. *zininzi* or *baninzi*).

Also relevant for this study is the use of loanwords in isiXhosa. Loanwords are words that are embraced by the speakers of one language (in this case isiXhosa) from another language (the source language) (O'Grady, Dobrovolsky, & Aronoff, 1997). In most cases nouns are borrowed; however, there are some languages that occasionally borrow verbs and adjectives (Brown, 2003). isiXhosa has many loanwords from English and Afrikaans, most of which are nouns (e.g. *ikati* 'cat'). However, isiXhosa also has a few verb stems that are loanwords from English or Afrikaans. These are used where no isiXhosa words are available and are used in a phonetically adapted form e.g. *-sarha* 'saw' (from Afrikaans 'saag') and *-bhaptiza* 'baptise' (from English) (Oosthuysen, 2016, p. 282).

In the following sections different formulations of isiXhosa comparative questions will be discussed, first in terms of

TABLE 4: Example of word order variation possible in isiXhosa.

isiXhosa	English
USigqibo unike uMveli isonka	<i>Sigqibo gave Mveli bread</i>
Umnike isonka uMveli uSigqibo	<i>Sigqibo gave Mveli bread</i>
Isonka uMveli usinike uSigqibo	<i>Sigqibo gave Mveli bread</i>
Usinike uMveli isonka uSigqibo	<i>Sigqibo gave Mveli bread</i>

Source: Oosthuysen, J.C. (2016). *The grammar of isiXhosa*. Stellenbosch: Sun Media

'classic' comparative questions and then in terms of 'alternative' comparative questions. Finally, a typology of isiXhosa compare type problems will be set up.

'Classic' comparative questions

Unlike English which only has one way to phrase the 'classic' comparative question, in isiXhosa there are a number of different ways in which the 'How many more' question can be phrased (see Figure 5). In this article three commonly used variations are discussed. One reason why variations are possible in isiXhosa is because there are two question words that can be used in combination with two words expressing 'more'. The two question words are *-ngaphi* 'how many' and *kangakanani* 'to what extent'. The two words used to express 'more' are the adjective *-ninzi* 'many/numerous/lots' and the adverb *ngaphezu(lu)* 'more/above' (see Mostert & Roberts, 2020, for a detailed discussion on the use of *-ninzi* and *ngaphezu(lu)* in comparison phrases).

The first formulation of a 'How many more' question in isiXhosa uses *-ngaphi* 'how many' and *ngaphezu(lu)* 'more/above'. In terms of word order, this formulation is the closest to the word order in English. Like in English, the question starts with 'how many' (*-ngaphi*). It is therefore possible that this formulation can result in a similar confusion as in English in that learners might answer the question *Zingaphi iimbiza?* 'How many pots?' instead of *Zingaphi iimbiza ngaphezu kweziciko?* 'How many more pots than lids?'

The second formulation also uses *-ngaphi* 'how many' but uses *-ninzi* 'many' to express 'more'. In this formulation, and in the third formulation, the adjective *-ninzi* 'many' is used before the question word *-ngaphi* 'how many'. Because the formulation does not start with *-ngaphi* 'how many', it is

Variation	Example
<u>-ngaphi?</u> + <u>ngaphezu(lu)</u>	Zingaphi iimbiza ngaphezulu kweziciko? zi-ngaphi iimbiza ngaphezulu kwe-ziciko they.are-how.many pots more compared.to-lids
<u>-ninzi</u> + <u>nga-</u> + <u>-ngaphi?</u>	Iimbiza zi-ninzi ngezingaphi kuneziciko?† iimbiza zi-ninzi nge-zi-ngaphi kune-ziciko pots they.are-many by-they.are-how.many compared.to-lids
<u>-ninzi</u> + <u>kanganani?</u>	Zininzi kangakanani iimbiza kuneziciko? zi-ninzi kangakanani iimbiza kune-ziciko they.are-many to.what.extent pots compared.to-lids

†, Alternatively: Zininzi ngezingaphi iimbiza kuneziciko? or Zininzi ngembiza ezingaphi kuneziciko?

FIGURE 5: Different formulations of the classic ‘how many more’ comparative question in isiXhosa.

possible that this formulation is less likely to result in learners answering the question ‘how many?’ instead of the question ‘how many more?’.

Note that due to the flexibility of word order in isiXhosa the second formulation can be expressed in a number of different ways. For this study, however, only one variation was considered (see footnote in Figure 5).

The third formulation uses *-ninzi* ‘many’ and a specialised question word *kanganani* ‘to what extent’. *Kanganani* can be used when asking about the difference between two nouns or sets of nouns. Because the use of *ngaphi*, it was speculated that this third formulation would be the least confusing for learners and therefore the easiest for them to solve. This was tested in this study by comparing the third formulation with one variation of the second formulation.

It is important to note that the question word *kanganani* ‘to what extent’ does not necessarily have to have a numerical answer. As in English, when asked, ‘How many more stars than squares?’, it is possible to answer ‘Many more’ or ‘A few more’. It is therefore important that isiXhosa learners, or at least teachers, are aware that the practice of answering a *kanganani* question with a numerical value is classroom based and is not necessarily used outside of the mathematics classroom.⁴

The canonical texts contain examples of all three formulations of classic comparative questions. While it is not possible to know what informed the choice of formulation in each example, this study sets out, in part, to provide research to better inform such decisions in the future.

Alternative comparative questions

As discussed previously, in English there are a number of alternative ways of asking comparative questions in conjunction with a matching problem situation. These include ‘how many won’t get’ (Hudson, 1980), ‘how many still needed’ (EGMA) and ‘how many missing’ (Roberts, 2016). Similarly it is possible to construct alternative comparative questions in isiXhosa, as exemplified in Figure 6.

4.Thank you to Bambilihle Nkwentsha for pointing this out.

Variation	Example
<u>-shota</u>	Ushota ngeziciko ezingaphi? u-shota nge-ziciko ezi-ngaphi she.is-short by-lids that.are-how.many She is short by how many lids?
<u>kusafuneka</u>	Kusafuneka iziciko ezingaphi? ku-sa-funeka iziciko ezi-ngaphi it.is-still-necessary lids that.are-how.many How many lids are still needed?
<u>ngazukufumana</u>	Zingaphi iimbiza ezingazukufumana iziciko? zi-ngaphi iimbiza ezi-nga-zu-ku-fumana iziciko they.are-how.many pots that.are-not-going-to-get lids How many pots won’t get a lid?

Note: Words defining the formulation are underlined.

FIGURE 6: Alternative comparative questions in isiXhosa (for matching (one-to-one embedded) problem situations).

While the *kusafuneka* ‘still need’ and the *ngazukufumana* ‘won’t get’ formulations have direct equivalents in English, the *-shota* formulation does not and therefore requires some additional comments. The adapted loanword *-shota* is a loanword from the English verb ‘be short of’. While in English ‘be short of’ is most commonly used to refer to money (e.g. ‘I am short (of) three rand’ to mean ‘I have three rand less than I need’), in isiXhosa *-shota* is commonly used to refer to being short of a wide range of things. In isiXhosa *-shota* is either used with the prefix *u-* when a person is short of something or with the prefix *ku-* when there is a shortage of things (not belonging to a specific person).

As part of a larger study, isiXhosa adults (both teachers and other caring adults) were observed engaging with isiXhosa learners and formulating comparative questions about specific problem situations. From these observations it appeared that the formulations that learners most easily understood were ones that included the verb *-shota*. This observation was the impetus for this study which, among other things, tests the hypothesis that *-shota* comparative questions are the easiest for learners to solve. Like the ‘how many missing?’ question introduced by Roberts (2016), asking ‘how many are short?’ draws attention to the absence of elements.

It is important to note that some isiXhosa speakers argue that it is not appropriate to use a loanword such as *-shota* in a mathematics classroom. Others argue that because it is a word learners are familiar with and understand, it should be used as a means to help learners make sense of compare type problems, at least in teaching, if not in formal testing.

Typology of isiXhosa compare type problems

Drawing on the typology of English comparative questions, and on the discussions of classic and alternative comparative questions in isiXhosa, Figure 7 provides a typology of isiXhosa compare (difference unknown) word problems. As with the English typology, a complete version of the typology with examples for each category is provided in Appendix 1, Figure 2-A1. The typology and the identification of the dimensions of variation (problem situation and comparative question) and the range of possible change for each dimension

Comparative question		Problem situation			
		Classic	No matching	Matching	
				1-to-1 embedded	1-to-1 not embedded
<i>ngaphezu(lu) + -ngaphi?</i>	more + how many?	-	-	-	
<i>-ninzi + nga + -ngaphi?</i>	many + by + how many?	Q1	Q3	-	
<i>-ninzi + kangakanani?</i>	many + to what extent?	Q4	-	-	
Alternative		No matching	1-to-1 embedded	1-to-1 imposed	
<i>ushota + -ngaphi?</i>	'short' + how many?	-	Q2	-	
<i>kusafuneka + -ngaphi?</i>	still necessary + how many?	n/a	-	Q5	
<i>ngazukufumana + -ngaphi?</i>	won't get + how many?	-	-	-	

n/a, not applicable.

FIGURE 7: Typology of isiXhosa compare type problems (with early grade mathematics assessment questions located in relevant categories).

make it possible to study the influence of one dimension by varying that dimension and keeping the other dimension the same.

Research design

In order to answer the two research questions, as set out in the introduction, results from the South African version of the EGMA, based on the core EGMA (Platas, Ketterlin-Geller, Brombacher, & Sitabkhan, 2014) and adapted by Brombacher and Associates, were used. The core EGMA includes one compare type problem out of a total of four additive relation word problems (Q5–Q8 in Table 5). For this study, four additional compare type word problems were also administered (Q1–Q4 in Table 5).

The four original word problems were translated from English into isiXhosa by an accredited translator. The additional four problems were formulated in isiXhosa through consultation with a number of isiXhosa speakers. See Appendix 1, Table 1-A1 for the isiXhosa formulation and English translations of the eight word problems.

Table 5 sets out the eight different types of word problems used in this study, based on the typologies set up in previous sections.

The research for this study was done in two stages, corresponding to the two research questions. In the first stage, two additional compare type problems (Q1 and Q2) were added to the EGMA assessment in order to answer the first research question, namely whether, in isiXhosa, different formulations of compare type problems had different levels of difficulty. This was answered by comparing Q1, Q2 and Q5. The results of this comparison (discussed below) confirmed that in isiXhosa, different formulations of compare type problems have different levels of difficulty. However, at this point it became apparent that the three formulations tested in the first stage differed in terms of more than one

TABLE 5: Different word problems used in this study.

Variable	Question	Problem type	Stage 1 (n = 242)	Stage 2 (n = 260)
Compare	Q1	No matching [- <i>ninzi</i> + - <i>ngaphi</i>]	x	-
	Q2	Matching (1-to-1 embedded) [- <i>shota</i>]	x	-
	Q3	Matching (1-to-1 embedded) [- <i>ninzi</i> + - <i>ngaphi</i>]	-	x
	Q4	No matching [- <i>ninzi</i> + <i>kangakanani</i>]	-	x
	Q5	Matching (1-to-1 imposed) [<i>kusafuneka</i>]	x	x
Other	Q6	Change (increase) – result unknown	x	x
	Q7	Change (increase) – start unknown	x	x
	Q8	Collection (attributes) – subset unknown	x	x

dimension. This led to stage two of the study in which the second research question was answered.

In stage two, in order to establish which of the different dimensions had an effect on the difficulty level, Q1 and Q2 were replaced with Q3 and Q4. The addition of Q3 and Q4 made it possible to isolate the effect of the comparative question (research question 2.1) by comparing two differently phrased questions with the same problem situation (Q1 and Q4 as well as Q2 and Q3). It also meant that it was possible to isolate the effect of the problem situation (research question 2.2) by comparing two problems with the comparative question formulated in the same way but with different problem situations (Q1 and Q3). The relationship between the four questions and where they are located in the typology of isiXhosa compare type problems is set out in Figure 8.

Methodology

In this section the methodologies used for data collection and data analysis are discussed in detail.

Data collection

The EGMA was administered to isiXhosa-speaking children in Grade 1, Grade 2 and Grade 3. The data were originally collected to evaluate an early grade mathematics intervention in five isiXhosa-dominant public schools in the rural Eastern

			Problem situation	
			No matching	Matching (1-to-1 embedded)
Comparative question	<i>-ninzi + nga + -ngaphi?</i>	many + by + how many?	Q1	Q3
	<i>-ninzi + kangakanani?</i>	many + to what extent?	Q4	-
	<i>-shota + -ngaphi?</i>	'short' + how many?	-	Q2

FIGURE 8: Questions used to isolate effect of different dimensions of variation.

Cape. All learners who were present on the day that the assessment was administered were tested.

The EGMA was administered twice in 2019, once in May (Stage 1, $n = 242$) and once in November (Stage 2, $n = 260$). Q1 and Q2 were added to the EGMA administered in May and Q3 and Q4 were added to the EGMA administered in November. Table 5 also indicates which questions were administered during each stage.

The EGMA was administered individually by isiXhosa-speaking adults. Each word problem was read to the learner in isiXhosa, first using isiXhosa number names, and then using English number names. Results (correct or incorrect) were recorded on tablets and then extracted into a spreadsheet for analysis.

The guidelines for administering the EGMA state that if a learner incorrectly answers four questions in a row, they should not be asked the remaining questions in that section, the assumption being that the learner would not be able to answer any of the remaining questions. For this article, in each stage, only the results of learners who were asked all six word problems were analysed.

Data analysis

In order to compare the difficulty level of two word problems, the facility score of each problem was calculated. The facility score is the percentage of learners who correctly answered the question out of the total number of learners who were asked the question. Questions with a higher facility score were considered to be easier than those with a lower facility score.

Because neither cohort answered all five compare type questions (see Table 5), it was necessary to establish whether the results of the questions administered only in Stage 1 could be compared with the questions administered only in Stage 2, and with those administered during both stages. In order to do this a Pearson's chi-squared test for homogeneity was done on facility scores of the four questions that were administered during both stages (Q5–Q8 in Table 5). The test returned a p -value of 0.96 indicating that the two groups of learners were very homogenous, in other words the learners

performed similarly on the four matched questions. In light of this, it is possible to compare any two questions, even if they were not answered by the same group of learners.

While early research on the relative difficulty of word problems only considered the facility scores of the problems, subsequent developments in data analysis techniques now allow for more sophisticated comparisons. In this study, for each research question the word problems were first compared in terms of their facility scores. If there was a difference in facility score, a Pearson's chi-squared test for independence was used to establish whether the difference in facility score was significant or not (see Appendix 1, Table 2-A1 for a summary of the p -values for each research question).

Ethical consideration

This study forms part of a PhD study which has received ethical approval from the University of Johannesburg, ethical clearance number: 2017-060. The data were anonymised and used only as aggregated data, which were not linked to individual children. Ethical clearance was received on 08 September 2017.

Results

The results will be discussed in relation to each research question. For each question the relevant problems will first be compared in terms of their facility scores and then in terms of the results of the Pearson's chi-squared test for independence.

RQ1: Relative difficulty of different isiXhosa compare type problems

In this section three different types of comparison problems are compared, namely a standard compare type problem (Q1), a matching (one-to-one embedded) problem (Q2), and a matching (one-to-one imposed) problem (Q5):

- (Q1) Iiswiti zentombazana zininzi ngeeswiti ezingaphi kwezenkwenkwe?
iiswiti ze-ntombazana zi-ninzi ngee-switi ezi-ngaphi
 sweets they.are-many in.terms. they.are-how.
 of-girl of-sweets many

kwe-ze-nkwenkwe

compared.to-of-boy

How many more sweets does the girl have than the boy?

(Q2) Ushota ngeziciko ezingaphi?

u-shota nge-ziciko ezi-ngaphi
she.is-short in.terms.of-lids they.are-how.many

She is short by how many lids?

(Q5) Kusafuneka iiofenji ezingaphi ukuze umama akwazi ukunika umntwana ngamnye iiofenji enye?

ku-sa-funeka iiofenji ezi-ngaphi ukuze umama a-kwazi
it.is-still- oranges that.are- so.that mother she-is.
necessary how.many able
uku-nika umntwana nga-mnye iiofenji e-nye
to-give child by-one orange that.is-one
How many oranges are still needed so that the mother can give each child one orange?

Figure 9 shows that, like in English, a smaller proportion of learners (47%) were able to solve standard compare type problems ('no matching' problem situation with a classic phrasing of the compare question) than both of the 'matching' problems (Q2 and Q5). There was less of a difference in the relative difficulty of the two different matching problems with a bigger proportion of learners able to answer the matching (one-to-one imposed) problem correctly (81%) than learners who answered the matching (one-to-one embedded) problem correctly (74%). The difference in facility scores was significant ($p < 0.05$).

RQ2: Effect of different factors on relative difficulty of word problems

In this section the different factors or dimensions that constitute a compare type problem are isolated to establish which factors influence the relative difficulty of compare type problems. The two factors that are considered are the problem situation and the phrasing of the comparative question. This is done by comparing different combinations of Q1–Q4, as outlined in the research design section.

Figure 10 shows that, for matching (one-to-one embedded) problems, a *-shota* formulation of the comparative question

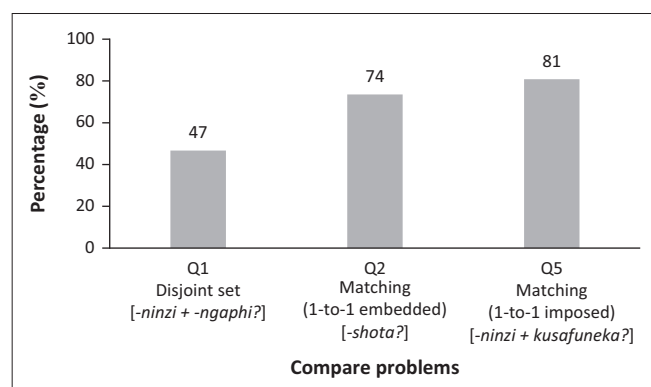


FIGURE 9: Facility score of three different isiXhosa compare type problems.

is easier for learners than a *-ninzi + -ngaphi* formulation. It also shows that, for no matching problems, a *kanganani* formulation is easier than a *-ninzi + -ngaphi* formulation. Finally, when a *-ninzi + -ngaphi* formulation is used, problems with a matching (one-to-one embedded) problem situation are easier than those with a 'no matching' problem situation. These results are discussed in detail in the following two sections.

RQ2.1: Effect of formulation of comparative question on relative difficulty

To investigate whether the formulation of the comparative question influences the relative difficulty of a word problem, two different problem types were considered: matching (one-to-one embedded) problems and standard compare type problems. For each problem type the same problem situation was used but two differently phrased questions were asked.

For matching (one-to-one embedded) problems, one question used the classic '*-ninzi + -ngaphi*' phrasing (Q3) and the other the alternative '*-shota*' phrasing (Q2):

- (Q3) Iimbiza zininzi ngezingaphi kuneziciko?
iimbiza zi-ninzi nge-zi-ngaphi kune-ziciko
pots they.are-many by-they.are-how.many compared.
with-pots
How many more pots than lids?
- (Q2) Ushota ngeziciko ezingaphi?
u-shota nge-ziciko e-zi-ngaphi
she.is-short in.terms.of-lids are-they-how.many
How many lids are short (missing)?

Figure 10 shows that fewer learners answered correctly when the '*-ninzi + -ngaphi*' phrasing was used (59%) and more learners answered correctly when the '*-shota*' phrasing was used (74%). The difference in facility scores is significant ($p < 0.05$). This comparison is similar to the comparison Hudson (1980) tested. Like with Hudson's comparison, it is possible that the alternative '*-shota*' phrasing is easier for learners to understand because it does not use quantifiers such as 'more' and 'less'.

For the 'no matching' compare problems, it was possible to compare two different formulations of the classic comparative

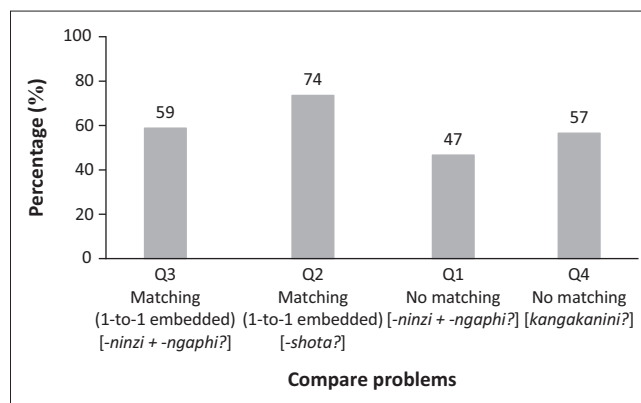


FIGURE 10: Facility scores of problems used to isolate factors influencing relative difficulty.

questions, namely the '*-ninzi + -zingaphi*' formulation (Q1) and the '*kangakanani*' formulation (Q4). As can be seen from the English translations, this comparison is not possible in English where there is only one formulation of the 'classic' comparative question.

(Q1) Iiswiti zentombazana zininzi ngeeswiti ezingaphi kwezenkwenkwe?

iiswiti ze-ntombazana zi-ninzi ngee-switi ezi-ngaphi

sweets of-girl they.are-many in.terms.of-sweets they.are-how.many

kwe-ze-nkwenkwe

compared.to-of-boy

How many more sweets does the girl have than the boy?

(Q4) Zininzi kangakanani iiswiti zentombazana kunezenkwenkwe?

zi-ninzi kangakanani iiswiti ze-ntombazana kune-ze-nkwenkwe?

they.are-many by.what.extent sweets of-girl compared.to-of-boy?

How many more sweets does the girl have than the boy?

Figure 10 shows that fewer learners answered correctly when the '*-ninzi + -ngaphi*' formulation was used (47%) and more answered correctly when the '*kangakanani*' formulation was used (57%). This supports the hypothesis that learners would find the '*kangakanani*' formulation less confusing than the '*-ninzi + -ngaphi*' formulation and begins to answer the question about which formulation of the 'classic' comparative question is most accessible to learners. Even though facility scores for these two problems are not that different, the difference is still significant ($p = 0.003 < 0.05$).

RQ2.2: Effect of problem situation on relative difficulty of word problem

In order to establish whether reframing the problem situation without changing the question also has an effect on the relative difficulty level of compare problems, the facility scores for a standard compare problem (Q1) and a matching (one-to-one embedded) problem (Q3), both with the same formulation of the 'classic' comparison question (*-ninzi + -ngaphi*), were compared:

(Q1) Iiswiti zentombazana zininzi ngeeswiti ezingaphi kwezenkwenkwe?

iiswiti ze-ntombazana zi-ninzi ngee-switi ezi-ngaphi

sweets of-girl they.are-many in.terms.of-sweets they.are-how.many

kwe-ze-nkwenkwe

compared.to-of-boy

How many more sweets does the girl have than the boy?

(Q3) Iimbiza zininzi ngezingaphi kuneziciko?

iimbiza zi-ninzi nge-zi-ngaphi kune-ziciko

pots they.are-many by-they.are-how.many compared.with-pots

How many more pots are there than lids?

Figure 10 shows that even when a more difficult 'classic' comparative question is used, a higher percentage of learners correctly answered the matching (one-to-one embedded) question (59%) than the percentage of learners who correctly answered the no matching problem (47%). This suggests that changing the problem situation and not the question can, on its own, make it easier for learners to understand the problem. Again, even though there is not a big difference between the facility score of these two questions, the chi-squared test ($p = 0.006 < 0.05$) confirms that the difference in facility score is significant.

Discussion

These results raise a number of points regarding the relative difficulty of isiXhosa compare type problems in early grade mathematics, some of which are also relevant for English. The study confirms that in isiXhosa, as in English, while standard compare type problems (no matching + classic comparative question) are the most difficult to solve, when a standard compare type problem is modified, either by changing the problem situation or the formulation of the comparative question, the problem can become significantly easier for learners to solve (see Figure 11).

The next two points are only relevant for isiXhosa as they relate to specialised words that do not have an English equivalent.

Classroom observations suggested that comparative questions using the loanword *-shota* (e.g. Q2) would be easier for learners to understand those using *-ninzi* and *-ngaphi* (e.g. Q3). This was confirmed by the results from this study.

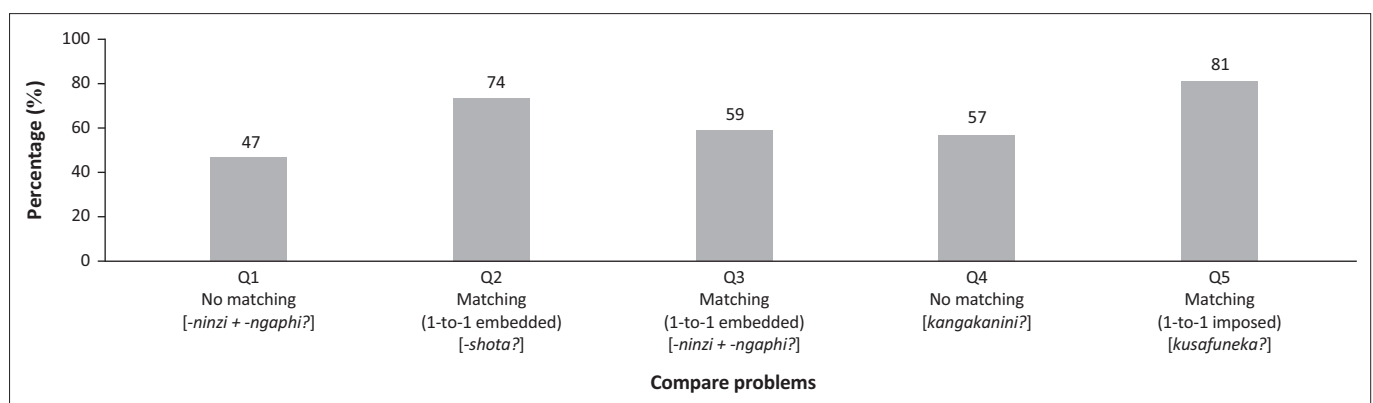


FIGURE 11: Facility score of different compare type word problems.

It is possible that this difference in difficulty level is because the *'-shota'* formulation does not use the quantifier 'more' while the *'-ninzi + -ngaphi'* formulation does. This difference in facility score suggests that teachers can use the *'-shota'* formulation to introduce learners to compare type problems.

It was speculated that the specialised question word *kangakanani* 'to what extent' would be less confusing for learners than questions using *-ninzi* and *-ngaphi* as these could be confused with *-ngaphi* questions. The results confirm this speculation: a higher percentage of learners correctly answered the *-kangakanani* question (Q4) than those that correctly answered the *'-ninzi + -ngaphi'* question (Q1) when the problem situation was kept the same. The fact that isiXhosa has a specialised question word that can be used when asking about the difference between two nouns or sets of nouns is an affordance that could be leveraged to mitigate the possible confusion between 'How many?' questions and 'How many more?' questions.

The final two points relate to issues that are also applicable in English.

While both matching (one-to-one embedded) and matching (one-to-one imposed) problem situations provide a useful teaching tool, in terms of the relative difficulty of comparison problems, matching (one-to-one imposed) problems are slightly easier to solve (see Figure 11). One possible explanation for this is that in matching (one-to-one imposed) problems (e.g. Q5) an additional phrase such as 'each child can get one orange' is required. This additional phrase makes the matching action explicit while in the matching (one-to-one embedded) problem (e.g. Q2), the matching action is implicit.

Finally, the influence of matching problem situations is not only observed when used together with alternative formulations of the comparative question. When a more difficult classic comparative question is used with both a 'no matching' problem situation (e.g. Q1) and with 'matching' problem situation (e.g. Q3), the problem with the 'matching' situation is still easier for learners to solve than the problem with the 'no matching' situation. It is possible that the reason for this is because the matching problem situation invokes the action of matching which can be used to solve the problem.

There are a number of limitations to be taken into account when interpreting the results. In order to fully establish whether there is a difference between matching (one-to-one embedded) and matching (one-to-one imposed) problems, it would be necessary to compare the two problem situations using all six different types of question. Similarly, in order to establish whether there is a difference between a *'-zingaphi'* question and a *'kangakanani'* question, it would be necessary to compare all three different compare type problems (no matching, matching (one-to-one embedded), matching (one-to-one imposed)). Finally, because the study did not compare all three classic comparative problems that can

be formulated in isiXhosa, it is not possible, at this stage, to establish which formulation is easiest for learners to understand.

Conclusion

Compare type word problems are notoriously difficult for learners but are also an important opportunity for learners to engage with the notion of comparison and of 'subtraction as difference'. While the 'standard' formulation of compare problems (no matching problem situation with classic comparative question) is difficult, this and other studies have shown that certain formulations of compare type problems are easier for learners to understand and to solve, both in English and in isiXhosa. These easier formulations provide a means of accessing the 'standard' compare type problems, allowing learners to make meaning of the problem situation without having to navigate complex language.

This article has contributed to understanding the different factors that constitute a compare type (difference unknown) word problem. The typologies of English and isiXhosa compare type problems provide a resource that can be used by materials developers and in further research looking in more detail at the influence of the different factors.

This article also highlights the importance of studying the ways in which African languages express mathematical ideas in order to identify and leverage affordances for teaching and learning mathematics and, where different formulations are possible, to establish which formulation is most accessible for learners. While this article lays a foundation for studying compare type problems in other Nguni languages, ultimately such research needs to be led by home language speakers in order for the linguistics features of African languages to be explored and described on their own terms, and not primarily in relation to English.

Acknowledgements

This study would not have been possible without the insight and guidance of Nicky Roberts and the generous contribution of many isiXhosa-speaking colleagues and friends, in particular: Yolisa Madolo, Nobuntu Mazeka, Zinyaswa Zuma, Bambelihle Nkwentsha, Zola Wababa, Tholisa Matheza, Hlumela Mkabile, Zikhona Gqibani, Nolutsha Parafini, Ezile Dalibango, Pumza Mfundisi and Thulelah Takane.

Competing interests

The author has declared that no competing interest exists.

Author's contributions

I declare that I am the sole author for this article.

Funding information

The study was funded by the South African National Research Foundation and the Department of Science and

Technology through Professor Henning, the South African Research Chair: Integrated Studies of Learning Language, Mathematics and Science in the Primary School, grant number 98573.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

Disclaimer

The views and opinions expressed in this article are those of the author and do not necessarily reflect the official policy or position of any affiliated agency of the author.

References

- Al-Murani, T. (2006). Teachers' awareness of dimensions of variation: A mathematics intervention project. *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education, 2*, 25–32.
- Al-Murani, T., Kilhamn, C., Morgan, D., & Watson, A. (2019). Opportunities for learning: The use of variation to analyse examples of a paradigm shift in teaching primary mathematics in England. *Research in Mathematics Education, 21*(1), 6–24. <https://doi.org/10.1080/14794802.2018.1511460>
- Berman, R.A. (2004). Between emergence and mastery: The long developmental route of language acquisition. In R.A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 9–34). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tilar.3.05ber>
- Brown, B. (2003). Code-convergent borrowing in Louisiana French. *Journal of Sociolinguistics, 7*(1), 3–23. <https://doi.org/10.1111/1467-9481.00208>
- Carpenter, T.P., Hiebert, J., & Moser, J.M. (1981). Problem structure and first-grade children's initial solution processes for simple addition and subtraction problems. *Journal for Research in Mathematics Education, 12*(1), 27–39. <https://doi.org/10.2307/748656>
- Carpenter, T.P., & Moser, J.M. (1983). The acquisition of addition and subtraction concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 7–44). New York, NY: Academic.
- Comrie, B., Haspelmath, M., & Bickel, B. (2008). *The Leipzig glossing rules: Conventions for inter-linear morpheme-by-morpheme glosses*. Retrieved from http://www.eva.mpg.de/lingua/pdf/LGR08_09_12.pdf
- Department of Basic Education. (2011). *Curriculum and Assessment Policy Standards (CAPS) grades 1–3: Mathematics. Curriculum document*. Pretoria: DBE.
- Department of Basic Education. (2012). *Diagnostic report: Annual national assessment 2012*. Pretoria: DBE.
- Department of Basic Education. (2014). *Annual National Assessment: 2013 Diagnostic report and 2014 framework for improvement*. Pretoria: DBE.
- Department of Basic Education. (2015). *Annual National Assessment 2014 diagnostic report foundation phase: Mathematics and home language*. Pretoria: DBE.
- Department of Basic Education. (2018). *Workbooks*. Retrieved from <https://www.education.gov.za/Curriculum/Workbooks/tabid/574/Default.aspx>
- Edmonds-Wathen, C. (2019). Linguistic methodologies for investigating and representing multiple languages in mathematics education research. *Research in Mathematics Education, 21*(2), 119–134. <https://doi.org/10.1080/14794802.2019.1615981>
- Fuson, K.C., Carroll, W.M., & Landis, J. (1996). Levels in conceptualizing and solving addition and subtraction compare word problems. *Cognition and Instruction, 14*(3), 345–371. https://doi.org/10.1207/s1532690xci1403_3
- Halliday, M.A.K. (1978). *An introduction to functional grammar*. London: Edward Arnold.
- Hudson, T. (1980). *Young children's difficulty with 'How many more __ than __ are there?' questions*. Bloomington, IN: Indiana University.
- Kennedy, C. (2009). Modes of comparison. In *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society*. Chicago, IL, 3–5 May.
- Lindvall, C.M., & Ibarra, C.G. (1980). Incorrect procedures used by primary grade pupils in solving open addition and subtraction sentences. *Journal for Research in Mathematics Education, 11*(1), 50–62. <https://doi.org/10.5951/jresmetheduc.11.1.0050>
- Marton, F., & Booth, S. (1997). *Learning and awareness*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mostert, I. (2019). Distribution of additive relation word problems in South African early grade Mathematics workbooks. *South African Journal of Childhood Education, 9*(1), 1–12. <https://doi.org/10.4102/sajce.v9i1.655>
- Mostert, I., & Roberts, N. (2020). Diversity of mathematical expression: The language of comparison in English and isiXhosa early grade mathematics texts. *Research in Mathematics Education*. Advance online publication. <https://doi.org/10.1080/14794802.2020.1821757>
- Nichols, J. (2007). What, if anything, is typology? *Linguistic Typology, 11*(1), 231–238. <https://doi.org/10.1515/LINGTY.2007.017>
- O'Grady, W., Dobrovolsky, M., & Aronoff, M. (Eds.). (1997). *Contemporary linguistics: An introduction* (3rd edn.). Boston, MA: St. Martin's Press.
- Oosthuysen, J.C. (2016). *The grammar of isiXhosa*. Stellenbosch: Sun Media.
- Petersen, B., McAuliffe, S., & Vermeulen, C. (2017). Writing and mathematical problem solving in Grade 3. *South African Journal of Childhood Education, 7*(1), 1–9. <https://doi.org/10.4102/sajce.v7i1.483>
- Platas, L.M., Ketterlin-Geller, L.R., Brombacher, A., & Sitabkhan, Y. (2014). *Early Grade Mathematics Assessment (EGMA) toolkit*. Retrieved from <https://www.edu-links.org/resources/early-grade-mathematics-assessment-egma-toolkit>
- Riley, M., Greeno, J., & Heller, J. (1983). Development of children's problem-solving ability in arithmetic. In H.P. Ginsburg (Ed.), *The development of Mathematical thinking* (pp. 153–196). New York, NY: Academic Press.
- Roberts, N. (2016). *Telling and illustrating additive relations stories: A classroom-based design experiment on young children's use of narrative in mathematics*. Johannesburg: University of Witwatersrand. <https://doi.org/10.13140/RG.2.1.1092.6964>
- Sepeng, P. (2013). Issues of language and Mathematics: Contexts and sense-making in word problem-solving. *Mediterranean Journal of Social Sciences, 4*(13), 51–62. <https://doi.org/10.5901/mjss.2013.v4n13p51>
- Spaull, N. (2016). Disentangling the language effect in South African schools: Measuring the impact of 'language of assessment' in grade 3 literacy and numeracy. *South African Journal of Childhood Education, 6*(1), 1–20. <https://doi.org/10.4102/sajce.v6i1.475>
- Verschaffel, L., & De Corte, E. (1993). A decade of research on word problem solving in Leuven: Theoretical, methodological, and practical outcomes. *Educational Psychology Review, 5*(3), 239–256. <https://doi.org/10.1007/BF01323046>
- Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity*. Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix starts on the next page →

Appendix 1

		Problem situation		
		No matching	Matching	
			1-to-1 embedded	1-to-1 not embedded
Comparative questions	Classic	A girl has seven sweets. A boy has five sweets.	A mother has eight pots and five lids.	A mother has seven children and two oranges.
	how many more?	How many more sweets does the girl have than the boy?	How many more pots than lids?	How many more children than oranges?
	Alternative	No matching	1-to-1 embedded	1-to-1 imposed
	how many missing?	-	How many lids are missing?	How many oranges are missing if the mother wants to give each child one orange?
	how many still needed?	n/a	How many lids are still needed?	How many oranges are still needed if the mother wants to give each child one orange?
how many won't get?	-	How many pots won't get a lid?	How many children won't get an orange if the mother gives each child one orange?	

Note: Grey highlighted text indicates phrase needed to impose one-to-one matching. n/a, not applicable.

FIGURE 1-A1: Typology of English compare word problems.

		Problem situation		
		No matching	Matching	
			1-to-1 embedded	1-to-1 not embedded
Comparative questions	Classic	Intombazana ineeswiti ezilithoba. Inkwenkwe ineeswiti ezine.	Umama uneembiza ezisibhozo neziciko ezihlanu.	Umama unabantwana abasixhenxe, abe neeorenji ezimbini.
	<i>ngaphezu(lu) + -ngaphi?</i>	Zingaphi iiswiti zentombazana ngaphezu kwe zenkwenkwe?	Zingaphi iimbiza ngaphezu kweziciko?	Bangaphi abantwana ngaphezu kweeorenji?
	<i>-ninzi + nga + -ngaphi?</i>	Iiswiti zentombazana zininzi ngeeswiti ezingaphi kwezenkwenkwe?†	limbiza zininzi ngezingaphi kuneziciko?†	Abantwana baninzi ngezingaphi kuneeorenji?
	<i>-ninzi + kangakanani?</i>	Zininzi kangakanani iiswiti zentombazana kunezenkwenkwe?†	Zininzi kangakanani iimbiza kuneziciko?	Baninzi kangakanani abantwana kuneeorenji?
	Alternative	No matching	1-to-1 embedded	1-to-1 imposed
<i>ushota + -ngaphi?</i>	-	Ushota ngeziciko ezingaphi?†	Ushota ngeorenji ezingaphi ukuze umama akwazi ukunika umntwana ngamnye iorenji enye?	
<i>kusafuneka + -ngaphi?</i>	n/a	Kusafuneka iziciko ezingaphi?	Kusafuneka iorenji ezingaphi ukuze umama akwazi ukunika umntwana ngamnye iorenji enye?†	
<i>ngazukufumana + -ngaphi?</i>	-	Zingaphi iimbiza ezingazukufumana iziciko?	Bangaphi abantwana abangazukufumana ukuba umama unika umntwana ngamnye iorenji enye?	

Note: †, indicates problems used in study. Grey indicates phrase needed to impose one-to-one matching. n/a, not applicable.

FIGURE 2-A1: Typology of isiXhosa compare word problems.

TABLE 1-A1: Compare problems used in study.

Question	Description
Q1	<i>Inkwenkwe ineeswiti ezine. Intombazana ineeswiti ezilithoba. Iiswiti zentombazana zininzi ngeeswiti ezingaphi kwezenkwenkwe?</i> A boy has four sweets. A girl has nine sweets. How many more sweets does the girl have than the boy?
Q2	<i>Umama uneembiza ezisibhozo neziciko ezihlanu. Ushota ngeziciko ezingaphi?</i> A mother has eight pots and five lids. How many lids is she short?
Q3	<i>Umama uneembiza ezisibhozo neziciko ezihlanu. Iimbiza zininzi ngezingaphi kuneziciko?</i> A mother has eight pots and five lids. How many more pots are there than lids?
Q4	<i>Inkwenkwe ineeswiti ezine. Intombazana ineeswiti ezilithoba. Zininzi kangakanani iiswiti zentombazana kunezenkwenkwe?</i> A boy has four sweets. A girl has nine sweets. To what extent are the girl's sweets more than the boy's sweets?
Q5	<i>Umama unabantwana abasixhenxe, abe neorenji ezimbini. Kusafuneka iorenji ezingaphi ukuze umama akwazi ukunika umntwana ngamnye iorenji enye?</i> A mother has seven children, and she has two oranges. How many oranges are still needed so that the mother can give the children each one orange?

TABLE 2-A1: *P*-values of problems that were compared for each research question.

RQ	Questions compared	Dimension compared	<i>p</i>
1	Q1, Q2, Q5	n/a	< 0.001*
2.1	Q1, Q4	different questions (same 'no matching' situation)	0.030*
2.1	Q2, Q3	different questions (same 'matching' situation)	< 0.001*
2.2	Q1, Q3	different problem situations (same '-ngaphi' question)	0.006*

n/a, not applicable.

*, *p* < 0.05 indicating significant differences in facility scores.