A brief critique of Hart, B. & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H. Brookes Publishing.

I.S.P. Nation
LALS, Victoria University of Wellington, New Zealand

Measuring vocabulary size from language output

The Hart and Risley (1995) study is an influential one. The book has had at least six printings, it is endorsed by leading educational figures, and its findings are widely cited. In essence, Hart and Risley's thesis is that the amount of language that young children hear affects their vocabulary size. Hart and Risley's three major conclusions (pages xix-xx) are (1) families differ greatly in the quantity of speaking that occurs with children, (2) quantity of speaking is linked to growth in vocabulary size, and (3) quantity of speaking is linked to the amount of quality features in the language used. The following diagram shows how the conclusions are related to each other.

Amount of speaking	→	Amount of quality features affecting vocabulary learning	→	Vocabulary size and growth
Conclusion 1		Conclusion 3		Conclusion 2

They also argue that because low socio-economic families speak less with their children, the vocabulary size of low socio-economic children is smaller than that of high socio-economic children. The more you hear, the more your vocabulary grows. Although this is an attractive and possibly true proposition, their research methodology is fundamentally flawed and so their study does not provide support for this proposition.

A cumulative count of types is not a measure of vocabulary size

The major flaw in the methodology is that Hart and Risley do not have independent measures for (1) the amount of language that children hear and (2) the vocabulary size of the children. In their methodology, the amount heard in interaction with parents and other family members and their measure of vocabulary size are effectively the same measure.

Hart and Risley's methodology was to regularly record (once a month) samples of speech each for the same length of time (one hour). For each child, the word types in each successive sample were added to an ever-increasing list. The size of the increasing list was considered to be a measure of the child's vocabulary size. The data was gathered in 42 families each month from when the child was 10 months old until they reached 36 months old.

We know from commonsense and from research data (Richards and Malvern, 1997) that as the number of tokens increase so will the number of types. The growth of the vocabulary list for each child is not "vocabulary growth" or "a developmental trajectory" but simply the result of an increasing accumulation of observations and thus an

increasing number of tokens and thus an increasing number of types. That is, the more tokens that you count, the greater the number of types you will find.

To make this clearer, let us take a slightly ridiculous example. Let us look at Shakespeare's plays beginning with his last play, and then moving play by play in reverse chronological order to the very first play he wrote. So, firstly we count the number of word types in his final play. Secondly, we count the number of word types in his second to last play. Many of the words in this play will also be in his final play, but there will be several words that only occur in his second to last play and these will be the words that we add to the list. By doing this for each play and adding the new word types to the list, we can show that Shakespeare's vocabulary grew from the writing of his last play to the writing of his first play -- which is of course ridiculous. Using the same methodology as Hart and Risley, I can show that my vocabulary size grew as I wrote each page of this report, doubling my vocabulary size by the end of the report.

	1st page	1st & 2nd pages	1st to 3rd pages	1st to 4th pages
Tokens	582	1222	1767	2276
Types	215	320	428	518

So, the major flaw in their methodology is the way they measured vocabulary size. It is possible that some of the new words produced were words that had been learnt in the previous month, but there is no way of knowing this from this measure.

Comparing the number of types in unequal samples is not comparing vocabulary sizes

The number of tokens in a sample directly affects the number of word types. When we compare the number of types produced, we must do this for equally sized samples of language. That is, the two samples must contain the same number of tokens (Richards and Malvern, 1997). If we do not do this, then the number of types will be strongly affected by the number of tokens. Hart and Risely compared the language produced by low socio-economic status families with the language produced by high socio-economic status families, and found that in the same time the low socio-economic status families produced less tokens. When the number of types in each group was compared, unsurprisingly there were more types produced in the high socio-economic status sample because of the greater number of tokens. Because number of types was considered to be a measure of vocabulary size, Hart and Risley concluded that children in lower socio-economic status families had slower vocabulary growth. This conclusion suffers from exactly the same fundamental flaw pointed out in the first section of this review. Comparing the number of types in differently sized samples of tokens is not comparing vocabulary sizes, it is comparing the effect of different numbers of tokens.

Limited samples of production do not show what children could do

There is a second major flaw. What someone produces is a very poor measure of their vocabulary size. There are many reasons for this. Firstly, we do not produce all the words

we know and could produce. There is a big difference between what we can do and what we actually do. Secondly, you would have to gather an enormous amount of output (tokens) to come anywhere approaching the number of types an individual knows. Even with young children, the number of tokens needed is likely to be in the thousands. Hart and Risley recorded children's output for an hour once a month. Thirdly, there are many factors besides vocabulary knowledge that can affect that amount of language that an individual produces. Some people are by nature quiet and thoughtful. Do they thus have a smaller vocabulary size than someone who talks endlessly? Fourthly, we also need to be aware that counting the words a child produces is a productive measure and their receptive vocabulary size will be much larger.

All these reasons why production is a poor measure of vocabulary size however, are a separate issue from the basic flaw in Hart and Risley's methodology – counting the increasing number of types in an increasing number of tokens is not measuring vocabulary size. It is simply measuring the effect of the number of tokens on the number of types.

Let us for a moment accept the idea that the number of types produced is a measure of vocabulary size. Because we know that as the number of tokens increases so does the number of types, to compare the vocabulary sizes of children we would need to compare the number of types *in exactly the same number of tokens* for each child (Richards and Malvern, 1997). If this is not done, then we are in fact not comparing vocabulary size but comparing the effect of different numbers of tokens. Because Hart and Risley's comparisons of highly socio-economic status and low socio-economic status children do not compare the number of types in the same number of tokens, they are not comparing vocabulary sizes. They are not making a valid comparison. They show that one group produces more tokens. They do not, however, show that one group has a larger vocabulary size.

Socio-economic differences in the quantity of language produced

Let us now look at a different issue – what might cause low socio-economic status families to produce less tokens in the study. Hart and Risley found that parents and children in low socio-economic status families produced fewer tokens in their one hour sessions than parents and children in high socio-economic status families. Because of this, the growth in the number of types was correspondingly less than in the high socio-economic status families. As we have seen above, the number of tokens and the number of types are inextricably related - less tokens results in less types.

Why did lower socio-economic families produce less data in interactions with their children? There are several possible reasons for this.

1 Lower socio-economic parents prefer to talk less.
2 Lower socio-economic families are much more guarded in their speech when they are being observed, and higher socio-economic families are much more forthcoming when they are being observed (see Labov (1969) for a discussion of

this). If this was the major reason, then the implication is that they would interact more when there wasn't an observer.

3 Lower socio-economic families may differ from higher socio-economic families on factors like the number of children in the family, the amount of work that a busy parent has to do, the number of other members of the family who are present, television watching habits and numerous other factors. Hart and Risley examined several of these factors and were confident that several of them could be eliminated as causative factors.

It is not clear from the study why the amount of interaction differed between the two social groups. It may be the case that this is a genuine difference between the two groups, although Labov (1969) would undoubtedly disagree.

Fortunately parent-child interaction is not the only source of vocabulary growth and it is highly likely that all native speakers will acquire a vocabulary size large enough to cope well with the everyday demands of working, education and social life. The question remains however what is the minimum amount of exposure to language which a young child needs to develop an adequate vocabulary size. It is clearly good that parents interact with their children to a large degree, but lower degrees of interaction may not necessarily result in an impoverished vocabulary. We need a properly-administered, independent measure of vocabulary size to determine this, and we need a broader picture of the sources of vocabulary learning.

Because of the flaw in the vocabulary size measure, the Hart and Risley study does not show that children in lower socio-economic families have smaller vocabulary sizes than children in higher socio-economic families.

Quantity and quality of language use

Hart and Risley found no major differences in the quality (richness) of the language that parents and children in low socio-economic status families produced. Hart and Risley looked at the quality of interaction in order to find examples of how interaction can help learning. If there is more of this helpful interaction then there will be presumably more learning. They note "All the quality features of language and interaction were present in the everyday parenting that all children experienced *but in differing amounts*" (p. 112, italics added). “the richness of quality features in parents’ utterances hardly varied. Whether a parent talked a lot or a little, such quality features as different words, questions, clauses, past-tense verbs, and affirmatives and prohibitions occurred approximately equally often per parent utterance.” (pp 127-8). Their argument about quality of interaction therefore still remains a quantity argument. That is, if lower socio-economic status parents did more of what they already do, it would be much better for the children.

Implications

I hope that the discussion in the first section of this critique has clearly shown that what has been used as a vocabulary size measurement in the study is not a vocabulary size measurement at all, but is simply the effect of the number of tokens on the number of types. It is very important to realise this, because Hart and Risley equate vocabulary size with cognitive functioning. "A vocabulary growth curve provides a direct and continuous measure of a child's intellectual functioning" (p. 16). If we accepted all of this, it would make the study a remarkable one and very self-contained, because in what is essentially one measure, we have measures of quantity of language produced, vocabulary size, and intellectual functioning.

The point of this critique is not to say that children's vocabularies don't grow. Clearly they do. The point is to show that the cumulative counting of word types in a series of limited language samples is not the way to measure vocabulary growth.

At best, the Hart and Risley study can say that low socio-economic parents speak less and interact less with their children than high socio-economic parents do, although the quality of the interaction does not differ. Because of the severe flaw in the vocabulary size measure, it is not possible to equate less interaction with lower vocabulary size.

The Hart and Risley study may be viewed as an attempt to explain why some children do not succeed at school. It does this by using a deficit model. In its crudest terms, this model says that low socio-economic status children do not get enough language-based interaction with their parents and as a result have a limited vocabulary. Because of this limited vocabulary, their cognitive functioning suffers. Because an adequate level of cognitive functioning is the basis of learning to read and succeeding in school subjects, low socio-economic status children do not do well at school. This vocabulary deficit limits them for the rest of their lives.

Because of the severe methodological flaws in the study, these conclusions are unwarranted. To truly investigate the relationship between quantity of interaction and vocabulary growth, we need at least two completely independent measures -- (1) a measure of quantity of interaction such as that used by Hart and Risley, and (2) a measure of vocabulary size such as a vocabulary size test.

References

Hart, B. & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H. Brookes Publishing.(2004 printing).

Labov, W. (1969). The logic of non-standard English. *Georgetown Monographs on Language and Linguistics 22*, 1-31. Reprinted in N. Keddie (ed.) (1973). *Tinker, Tailor … The myth of cultural deprivation*. (pp. 21-66) Harmondsworth: Penguin.

Richards, B. J. & Malvern, D. D. (1997). *Quantifying lexical diversity in the study of language development*. The New Bulmershe Papers. Reading: University of Reading.