

MLU AND IPSYN MEASURING ABSOLUTE COMPLEXITY

Lea Nieminen

Abstract. This article compares the results of Mean Length of Utterance (MLU) and Index of Productive Syntax (IPSyn) with the structural complexity of spontaneous utterances produced by 30-month-old Finnish children in a semi-structured playing situation. The comparison was carried out in order to determine the aspects of structural complexity which can be detected with MLU and IPSyn. This research adopts the frameworks of absolute complexity together with a multidimensional view of utterance structure and, furthermore, applies it through Utterance Analysis (UA). The results of the comparison between the metrics and changes in structural complexity discovered by UA reveal that MLU and IPSyn do function as measures of structural complexity but only if used in close relation to each other. Because they focus on different aspects of utterances, the results of both metrics should be interpreted in relation to one another.

Keywords: morphosyntax, acquisition, structural complexity, child language, Finnish

Introduction

This article focuses on the morphosyntactic complexity of spontaneous speech of children and on the methods used to evaluate structural complexity. The study aims to find out how structural complexity grows in children's utterances and how the existing metrics that have been developed to measure this react to changes in complexity.

The study was originally inspired by two crucial facts. Firstly, complexity is a frequently used notion in child language studies and it is often connected to development. The growth in complexity is perceived as an index of development in linguistic abilities and, therefore, complexity is an essential concept in language acquisition studies. However, despite its essential role, the concept of complexity has been used in acquisition studies without being properly defined, and this is the

second fact that has led to the current frame of research. The lack of an adequate definition often leads to a situation in which it is not clear what the results actually mean and which properties of language they do and do not describe.

In this study, two well-known complexity metrics, Mean Length of Utterance (MLU) and Index of Productive Syntax (IPSyn), are linked to an absolute approach to complexity. The main research questions are whether these metrics react to the growth of absolute complexity and whether the ways in which the metrics react to it are similar or dissimilar.

Linguistic complexity is a complicated concept that can be defined in several different ways. Chapter 2 introduces two separate perspectives of the concept, while Chapter 3 provides a short summary of how complexity has been traditionally discussed in language acquisition studies, and Chapter 4 focuses more on MLU and IPSyn, which are the most common metrics in evaluations of complexity. Chapters 5, 6 and 7 describe the current study in detail, and Chapter 8 concludes the paper with a discussion of the most important findings.

Multifaceted complexity

Complexity has been and may still be a somewhat dangerous notion in the field of linguistics, because it is easy to connect the complexity or simplicity of language to the cognitive abilities of the language users in an over-simplified manner (Kusters 2003: 2). There is also a well known claim that all languages are equal in terms of their complexity (Hockett 1958). This theory suggests that if a given language has a more complex morphology than another, this is offset by, for example, a simpler syntax, meaning that the overall complexity of the given language is kept in balance with all other languages. These are probably the main reasons why more detailed definitions and discussions regarding linguistic complexity have not been addressed until the last few years. As a consequence of recent discussions, especially among language typologists (e.g. McWhorther 2001, Kusters 2003, Dahl 2004 and Miestamo 2006) the multifaceted nature of linguistic complexity has become an important issue.

Linguistic complexity can be viewed as a relationship between a language user and the language. According to this so-called relative approach to complexity, it is crucial to emphasise who perceives a given language to be complex (Kusters 2003: 6), where complexity is defined as the level of difficulty or ease experienced by a language user. It is clear, for example, that a native speaker and a language learner will evaluate the complexity of the same language or linguistic structure differently, and that a person's evaluations may vary over time.

In connection to relative approach, answering questions such as "What causes the difficulty?" or "Why is X easier than Z?" requires a user-based perspective. When using language, a language user is involved in several processes, both for receiving and producing language. Are some linguistic structures difficult to perceive but easy to produce, or does the complexity become apparent only when pronouncing the structure? Are the linguistic means relatively easy, while the understanding of the ultimate meaning of the structure places extra pressure on cognitive processes? Is reading easier than listening or vice versa? Do all linguistic processes have equal weight when language users attempt to analyse their experiences of difficulty? These

questions, among others, highlight the need to discuss as many factors as possible in connection with complexity evaluations from a relative point of view.

Another approach to linguistic complexity focuses on the language itself and excludes the language users. This so-called absolute view approaches linguistic complexity from an information theoretical perspective, viewing complexity as a measurable object and defining it as the length of the shortest specification or description of the object. The notion of complexity should be kept distinct from the use and the users of the system whose complexity is evaluated, in order to maintain objectivity. (Dahl 2004: 21, 39–40) The absolute approach concentrates on counting the number of linguistic units, elaborations, rules, and so on (Miestamo 2006). The more such counted units there are, the more complex the language or evaluated linguistic structure.

Three important remarks must be made in connection with the absolute approach to linguistic complexity. Firstly, the definition of a language or a linguistic structure as being more complex than another does not constitute proof that that language is more difficult than any other. In other words, complexity and difficulty are not the same thing in the absolute approach. Secondly, a simpler grammar does not necessarily mean that a language is somehow primitive or inefficient as a means of communication (Miestamo 2006, McWorther 2001), and therefore the evaluation of complexity is by no means a value judgement. Thirdly, the absolute approach is, in a way, relative in nature because the demarcation of units to be counted depends on the linguistic theory that is represented or applied by the evaluator. As Kusters (2003: 6) has claimed, this approach is logically impossible if “absolute” is interpreted literally. It may be less provocative, therefore, to describe the absolute approach as a language-based approach to complexity as a contrast to the relative, user-based view of complexity (Nieminen 2007: 34).

Complexity in child language studies

Complexity is a commonly used notion in child language acquisition studies. In this particular field of linguistics, complexity has not been considered as a dangerous or provocative notion. On the contrary, the growth of complexity is synonymous with linguistic development; more advanced is also more complex. Comparing children in terms of their linguistic development is akin to comparing their positions on a scale of complexity. Acquisition studies have assumed that children acquire simple structures first and then more complex structures (Smith, van Kleeck 1986, Smith 1988), so the acquisition order would also reveal the complexity order. Whether this order is determined by the relative or absolute complexity of the items (or both) is not clear. The concept of complexity is usually taken for granted and neither proper identification nor thorough consideration of this multifaceted notion has been covered in detail (Nieminen 2007: 17).

Significant issues, such as considering linguistic or cognitive complexity, as well as their mutual relationship (cf. Bowerman 1996), have tended to be left to one side. Furthermore, only a few researchers have considered the role of various linguistic processes in complexity: whether linguistic items are difficult to perceive, acquire, produce or understand and how these processes affect the overall complexity of an item (cf. Peters 1997). As stated above, the complexity of a given linguistic item

may originate from various sources, and researchers should carefully define which of these factors is the focus of their studies.

Even in acquisition studies that clearly focus on linguistic structure, the demarcation of complexity has been implicit, and what is considered as complex has usually become evident only after interpretation of the methods and the results. The definition of complexity has thus far been conducted more comprehensively in disfluency studies, which are confined to dealing with, for example, sentence structure (e.g. Ratner, Sih 1987, Yaruss 1999). A more careful and limited definition of complexity has been found to be essential, due to findings claiming a strong connections between disfluencies and the structural complexity of utterances (see Ratner 1995, and Silverman, Ratner 1997, for review).

Complexity metrics and utterance structure

One thing that acquisition studies present clearly is that complexity is a measurable property of language, and different scales of measurement are used to execute this. The best known and most widely used measure for this purpose is Mean Length of Utterance (MLU), which was first introduced in its present form by Roger Brown ([1973] 1976). Another commonly used measure is the Index of Productive Syntax (IPSyn; Scarborough 1990a).

The basic premise of MLU is very simple: the number of productive morphemes produced in 100 consecutive intelligible utterances is calculated and divided by the number of utterances to determine the average number of morphemes per utterance. MLU is based on the idea that “almost every new kind of knowledge increases length: the number of semantic roles expressed in a sentence, the addition of obligatory morphemes, coding modulations of meaning, the addition of negative forms and auxiliaries used in interrogative and negative modalities, and of course, imbedding and coordinating” (Brown [1973] 1976: 77). MLU calculations can vary in many ways, and the results are not unambiguously cross-linguistically comparable. Also, its relation to complexity has been formulated in various expressions. MLU has been described as “a developmental index of language proficiency” (Johnston 2001: 161), “a general indicator of structural development” (Miller 1981: 25), “a valid predictor of syntactic complexity and diversity” (Rondal et al. 1987: 444) and an “index of clausal complexity” (Blake et al. 1993: 139).

IPSyn is based on morphological forms and syntactic structures that have been found to be essential in child language. The selection of predefined structures is always language specific, so the results of IPSyn are not cross-linguistically comparable. The IPSyn score is determined based on the number of first occurrences of the predefined structures in a child’s spontaneous speech and possible second occurrences in a different linguistic context. The ranking of the structures is usually done for 100 intelligible spontaneous utterances (for detailed instructions, see Scarborough 1990a, and Nieminen, Torvelainen 2003).

MLU and IPSyn have several similarities. Firstly, they both represent an absolute approach to complexity, since the complexity of a child’s utterances or their linguistic repertoire is determined by the number of calculated units, that is, morphemes in MLU and predefined structures in IPSyn (Nieminen 2007: 56). Secondly, they are both used to analyse spontaneous speech, and thirdly, the section of speech

to be analysed in both methods is an utterance, which is considered to be the most fundamental psycholinguistic unit (Tomasello 2000: 63). Finally, they both focus on morphosyntactic structure of utterances. Previous studies, such as Scarborough et al. (1991), found a high positive correlation between the two measures. It has also been argued that they both measure the same properties of language, so they have been used in studies to validate one another (e.g. Scarborough 1990b). Despite the high positive correlation, however, Nieminen (2007) presents opposing conclusions concerning the properties of language that are actually measured with these metrics. She claims that MLU offers a general view of utterances as a whole, although this is done by flattening the utterance structure into a string of morphemes. In contrast, IPSyn is a more analytic metric in that it attempts to extract the different morphosyntactic structures used to build the utterances and, therefore, it also deals with relations between linguistic elements. However, the overall structure of utterances is set aside when focusing on separate building blocks. Probably the most striking difference between these two metrics is the final targets of their evaluation. MLU evaluates actual speech production, while the target of IPSyn is the size and quality of a child's morphosyntactic repertoire. (Nieminen 2007: 58–59)

Both MLU and IPSyn are used to analyse the complexity of spontaneous speech. According to Nieminen (2007), utterance structure should be understood as a multidimensional whole consisting of individual syntactic components that may be composed of several morphological and syntactical elaboration layers. The structural complexity of an utterance therefore emerges not only from the number of components but also from their layered inner structures. Thus, complexity can grow in at least two ways: by adding new components or by adding new layers to existing components (Nieminen 2007: 39). Acceptance of two different dimensions for the growth of complexity raises the question of how MLU and IPSyn react to this, because they each seem capable of noticing only one kind of structural change. In other words, the metrics imply that complexity can only grow in one way, although the utterance structure offers several possibilities for complexity to change. This study was conducted in order to investigate this clear discrepancy. Its primary area of interest is the aspects of absolute complexity recognised by MLU and IPSyn and how the metrics work together in analysing the speech production of young children.

Subjects, data and methods

The participants of the study were 40 normally developing 30-month-old children (22 boys and 18 girls) acquiring Finnish as their first language. The data was recorded at the children's homes, in a semi-structured playing situation in the presence of an adult. All children were given the same toys to play with and each recording lasted for approximately 20 minutes. The recordings were transcribed and the children's productions were coded morphologically using CHILDES (MacWhinney 1991).

The MLU programme provided by CHILDES was used for the MLU calculation, which was based on the 80 longest utterances produced by each child. The 80 longest utterances were chosen instead of 100 consecutive intelligible utterances to ensure that each child's production skills were represented by the data sample in the most versatile manner. The same utterances were also analysed using the Finnish ver-

sion of IPSyn, which focuses on 49 morphological, syntactic and morphosyntactic structures in the Finnish language (Nieminen, Torvelainen 2003).

The basic purpose of this study was to test what kind of features of structural complexity these metrics could detect and react to. Therefore, the results from MLU and IPSyn measures needed to be contrasted with a detailed analysis of structural complexity, so a new analysis method was created for this purpose. The Utterance Analysis (UA) provides a detailed description of syntactic components and layers in each utterance. Because of the multidimensional approach to utterance structure it provides a comprehensive picture of the structural features that are involved in the composition of complexity.¹

MLU and IPSyn results

The 80 longest utterances in terms of the number of morphemes were extracted from each child's productions and analysed. Both MLU and IPSyn showed a wide variation in the data sample, as shown in Figure 1. The MLU values ranged from 1.233 to 7.862, whereas the IPSyn scores varied between 10 and 86 points. This is not surprising since, at the age of 30 months, children usually have widely varying levels of speech production. However, the high MLU values are particularly striking, but this is explained by the selection criteria. When only the longest utterances are included, all or most of the one-morphemic and even two-morphemic utterances are omitted and therefore do not decrease the average values.

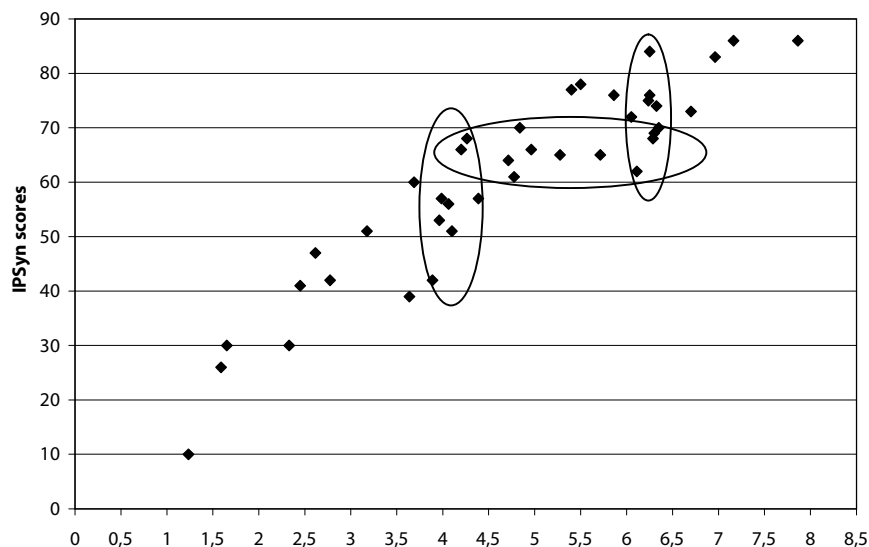


Figure 1. The relationship between MLU values and IPSyn scores

One way to evaluate the relationship between MLU values and IPSyn scores is to calculate the correlation between the results. The correlation was found to be very high in this study, ($r = .92$), which is in keeping with the findings of Scarborough et

¹ For a more detailed description of UA, see Nieminen 2007: 86–99.

al. (1991). However, a closer look at the individual scores presents a different picture of the relationship. There are points at which IPSyn scores imply great differences between children but MLU remains in a very narrow range, and vice versa. In other words, some linguistic features seem to be detected and interpreted as development only by one of the metrics, while the other metric does not react to these changes or reacts only very mildly. These contradictory results are circled in Figure 1.

The first of these contradictory points is between the MLU values 3.688 and 4.388. On the MLU scale the change is very small but in IPSyn scores the variation is as wide as from 42 to 68 points, covering more than 25% of the whole IPSyn scale (the maximum number of points is 98). In practice this implies that the utterance length is approximately four morphemes but different children create these four-morphemic utterances in significantly different ways. The same kind of discrepancy between the MLU and the IPSyn can also be seen in the upper part of the figure: the MLU stays constantly between 6.050 and 6.350 but the IPSyn scores vary between 62 and 84 points. A contrasting case of discrepancy can be found between the cases mentioned above. While the MLU values vary from 3.688 up to 6.350, the IPSyn scores remain between 60 and 70 points. In other words, while the length of utterances grows, the repertoire of structures used to build them remains at a constant level. This is possible only if the structures that first occur alone in utterances are gradually combined to create longer productions of speech.

Detailed analysis of structural complexity – contradictory MLU values and IPSyn scores against Utterance Analysis

Utterance Analysis describes in detail the kind of units that an utterance consists of and the relationship between the units. It provides a picture of the number of syntactic components (NP, V, AP, PP, CONJ etc.) involved in an utterance and what kind of elaboration (inflection, use of determiners or qualifiers) is present in each component (see Figure 2 for an example of Utterance Analysis).

1.	ADVP ADV:INT	V		NP
2.	INES		PRON:DEM	N
3.			GEN	
	<i>Missä</i>	<i>on</i>	<i>sel</i>	<i>lautanen?</i>
	Mi-ssä	on	se-n	lautanen?
	Wh-INES	is	it-GEN	plate?
	'Where is its plate?'			

Figure 2. An example of Utterance Analysis

When the data samples presenting contradictory MLU and IPSyn results were subjected to more detailed analysis through Utterance Analysis, different strategies and unexpected similarities were found in the creation of structural complexity. The first survey revealed that the number and type of syntactic components that the children used in their utterances were very similar, although the varia-

tion in MLU and IPSyn results seemed to suggest that the utterances were quite different. Approximately 90% of all utterances consisted of two, three or four syntactic components and approximately 90% of these components represented nominal phrases, verb constructions and adverb phrases. This suggests that the basic component structure of utterances is created during an early developmental phase and that the growth of structural complexity is mainly created by syntactic and morphological elaboration of the components. However, even in elaboration of utterances the children who were first found by MLU and IPSyn to be at different developmental levels showed remarkable similarities, especially when their utterances contained more than two syntactic components. Thus, the distinctive features of elaboration and structural complexity are clustered, especially in two-component utterances.

The first contradictory group of children had the MLU values ranging between 3.688 and 4.388 but the IPSyn scores showed remarkable variation, between 42 and 68 points. This group of children used a very limited repertoire of elaborations in their utterances. The limitations were two-fold in nature. Firstly, in most cases only one syntactic component of an utterance was elaborated while the others were left unelaborated. Secondly, these children almost exclusively used morphological elaboration, that is, inflection, and modifiers or other kind of syntactic elaborations remained quite rare. The variation in IPSyn scores can be explained by the fact that these children seem to be in a rapid phase of morphological acquisition. Their morphological repertoires are remarkably different but all of the children use their repertoires in a similar way, no matter how large or narrow it is.

The second contradictory group represents the children with similar IPSyn scores (60–70 points) together with a wide range of MLU values (3.688–6.350). Utterance Analysis revealed that, for these children, it is characteristic to expand elaboration to several syntactic components in utterances, and children do not need to apply many new linguistic means in order to achieve this. At the same time the elaboration is still mainly morphological, although syntactic elaboration does increase its proportion, and the elaboration does not, therefore, create many “new” morphosyntactic structures that could be recognised by IPSyn either. These factors explain why IPSyn scores remain on a constant level. The growing MLU is explained by the increased use of morphological elaboration, the spreading of elaboration from one to several syntactic components, and the combination of several resources in one utterance but not yet within one syntactic component. Structural complexity is grown by means that were acquired earlier and are now gradually put into full use.

The third group of contradictory results represented a similar pattern to that of the first group. MLU exhibited only very small changes (6.050–6.350) whereas the IPSyn varied between 62 and 84 points. Utterance Analyses revealed that these children constantly combined individual resources, not only within an utterance but also within a syntactic component. In other words, these children elaborated syntactic components by creating more elaboration than children in the other two groups. Both morphological and syntactic elaborations are used, and they occur together in the same syntactic components, a combination of linguistic resources that also creates structures that are recognised by IPSyn as “new”. Naturally, this raises the number of IPSyn points, but the length of utterances does not necessarily grow because the average of six morphemes already provides space for many kinds of expressions and linguistic variations. Earlier, structural complexity had spread

into more components but now expansion occurs simultaneously in more components and also adds new layers to the components.

Comparing MLU values and IPSyn scores with Utterance Analysis results reveals several interesting aspects of structural complexity and how it is recognised by the commonly used complexity metrics. Both MLU and IPSyn results show growth as well as steady phases, but the metrics do this crosswise. Utterance Analysis, however, reveals that although one metric or the other indicates a steady phase, the structural complexity of utterances is in a state of change. This strongly suggests that neither of the metrics can recognise all means of complexity growth. A more realistic picture of the changes in structural complexity can only be gained if MLU and IPSyn are used together in close mutual collaboration and are interpreted in relation to each other. Interpretation of the results requires resources to be in proportion to the average length of utterances.

Discussion

The goal of this research was to test how well the commonly used complexity metrics, MLU and IPSyn, carry out the basic task for which they were originally invented. Is it possible for them to recognise the growth of complexity? What aspects of absolute complexity do they react to? Do both of the metrics react to the same aspects or do they differ in terms of the type of language properties they focus on? Are these metrics compatible with a multidimensional view of utterance structure?

MLU in particular has been a target for wide and varying criticism (for a review, see Rollins et al. 1996). However, this study indicates that a simple measure of average utterance length can catch some valuable aspects of growing structural complexity. When children's MLU values were close to 4.000, the structural complexity of their utterances did not change very much; instead, it was the repertoire of structural resources that varied. MLU was also able to detect the starting point of the combining of resources, because this process also automatically increased the length of utterances. The resource combinations also increase the structural complexity at both the component and utterance levels. However, MLU stopped reacting to the combining of resources once it was spread wider. The utterances are now so long that there is ample space for a range of structure combinations and complexity levels without a need to expand the utterance length. At this point of development, resource combinations create "new" structures from IPSyn's perspective. IPSyn is a cumulative metric (Nieminen, Torvelainen 2003, Nieminen 2007: 54) and therefore it observes growth within an individual structure and also gives credit for it.

Although MLU was found to probably be a better detector of growth in structural complexity than the former criticism would suggest, the interpretation of MLU values could not be possible without the help of IPSyn scores. If IPSyn results were not available it would not have been possible to conclude that utterance structures do not vary a lot from child to child when MLU is approximately 4.000 but that children's linguistic skills may vary considerably according to the resources they have. In addition, it is impossible to say that growth of MLU from approximately 4.000 to 6.000 is due to a combination of structures in utterances with the corresponding IPSyn scores that stay at a constant level. These kinds of interpretations

suggest that neither of the metrics can be used as a reliable index of structural complexity without the other. A valid evaluation of complexity requires the use of both metrics and, in particular, synchronous interpretation when the final results are concluded.

The patterns of structural development that were discovered in this study should provide food for thought for those who have used MLU as a selection device. In many studies, for example, normally developing children and language-impaired children have been matched according to their MLU values, because similar MLUs have been connected to similar skills in expressive language. However, the results of this study indicate that the concepts of similarity must be considered with caution and that one metric does not equate to a thorough investigation in this respect.

A secondary result of this study was the discovery of developmental trends in growth of structural complexity. The starting point was that utterance structures are multidimensional rather than linear, and structural complexity may therefore grow in several dimensions. The Utterance Analysis results demonstrated that Finnish children start with morphological elaboration, in other words they use inflectional forms of nouns and verbs. Whether these forms are memorised as a whole or are actually inflected by a child remains unclear, but the fact that children differed substantially in their linguistic resources but not in their use of these resources might suggest that at this point of development it is more a lexical than morphological phenomenon. In other words, the inflectional forms are likely to be rote learnt forms. Unanalysed chunks and frame patterns have long been strongly supported as candidates for units of acquisition (see, for example, Peters 1983, Kauppinen 1998, Tomasello 2003). The next developmental phase, which involved adding elaboration to several syntactic components, may still basically be the addition of rote learnt units, but more research would be needed for this to be proven. The last development phase detected in this study was that at which elaboration had spread into several syntactic components of an utterance and children also combined different elaborations within a component, thereby creating multilayered structures. In particular, the combining of elaboration resources within a syntactic component suggests that these children already use analytic processing of smaller linguistic units. Combining elements requires that these elements have been extracted from the units where they were originally acquired.

This research was initially conducted in order to shed light on the concept of complexity and to find out what is actually being investigated when so-called complexity metrics are used as an analysing method. The study's results have shown that MLU and IPSyn function as spotlights that shed light on children's expressive language from different perspectives. The best and most reliable results can be achieved when both metrics are used simultaneously and interpreted in close relation to one another.

References

- Blake, Joanna; Quartaro, Georgia; Onorati, Susan 1993. Evaluating quantitative measure of grammatical complexity in spontaneous speech samples. – *Journal of Child Language*, 20, 139–152. doi:10.1017/S0305000900009168
- Bowerman, Melissa 1996. The origin of children's spatial semantic categories: Cognitive versus linguistic determinants. – John J. Gumperz, Stephen C. Levinson (Eds.). *Rethinking*

- Linguistic Relativity. *Studies in the Social and Cultural Foundations of Language* 17. Cambridge: Cambridge University Press, 145–176. doi:10.2277/0521448905
- Brown, Roger [1973] 1976. *A First Language. The Early Stages*. Harmondsworth: Penguin Education.
- Dahl, Östen 2004. *The Growth and Maintenance of Linguistic Complexity*. *Studies in Language Companion Series* 71. Amsterdam: John Benjamins.
- Hockett, Charles B. 1958. *A Course in Modern Linguistics*. New York: Macmillan.
- Johnston, Judith R. 2001. An alternate MLU calculation: Magnitude and variability of effects. – *Journal of Speech, Language, and Hearing Research*, 44, 156–164. doi:10.1044/1092-4388(2001/014)
- Kauppinen, Anneli 1998. Puhekuviot, tilanteen ja rakenteen liitto. Tutkimus kielen omaksumisesta ja suomen konditionaalista. *Suomalaisen kirjallisuuden seuran toimituksia* 713. Helsinki: Suomalaisen kirjallisuuden seura.
- Kusters, Wouter 2003. *Linguistic Complexity. The Influence of Social Change on Verbal Inflection*. LOT Dissertation Series 77. University of Leiden. Utrecht: LOT.
- MacWhinney, Brian 1991. *The CHILDES Project: Tools for Analyzing Talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McWhorter, John H. 2001. The world's simplest grammars are creole grammars. – *Linguistic Typology*, 5, 125–166. doi:10.1515/lity.2001.001
- Miestamo, Matti 2006. On the feasibility of complexity metrics. – Krista Kerge, Maria-Maren Sepper (Eds.). *FinEst Linguistics. Proceedings of the Annual Finnish and Estonian Conference of Linguistics*. Tallinn, May 6–7, 2004. Tallinn: Tallinn University Press, 11–26.
- Miller, Jon F. 1981. *Assessing Language Production in Children. Experimental Procedures*. London: Edward Arnold.
- Nieminen, Lea 2007. *A Complex Case: A Morphosyntactic Approach to Complexity in Early Child Language*. *Jyväskylä Studies in Humanities* 72. Jyväskylä: University of Jyväskylä.
- Nieminen, Lea; Torvelainen, Päivi 2003. Produktiivisen syntaksin indeksi – suomenkielinen versio. – *Puhe ja kieli*, 23, 119–132.
- Peters, Ann M. 1983. *The Units of Language Acquisition*. *Cambridge Monographs and Texts in Applied Psycholinguistics*. Cambridge: Cambridge University press.
- Peters, Ann M. 1997. Language typology, prosody, and the acquisition of grammatical morphemes. – Dan I. Slobin (Ed.). *The Crosslinguistic Study of Language Acquisition*. Volume 5: Expanding the Context. Mahwah, NJ: Lawrence Erlbaum Associates, 135–197.
- Ratner, Nan Bernstein 1995. Language complexity and stuttering in children. – *Topics in Language Disorders*, 15 (3), 32–47.
- Ratner, Nan Bernstein; Sih, Catherine C. 1987. Effects of gradual increases in sentence length and complexity on children's disfluency. – *Journal of Speech and Hearing Disorders*, 52, 278–287.
- Rondal, Jean A., Ghiotto, Martine; Brédart, Serge; Bachelet, Jean-François 1987. Age-relation, reliability and grammatical validity of measures of utterance length. – *Journal of Child Language*, 14, 433–446. doi:10.1017/S0305000900010229
- Rollins, Pamela R.; Snow, Catherine E.; Willet, John B. 1996. Predictors of MLU: Semantic and morphological developments. – *First Language*, 16, 243–259. doi:10.1177/014272379601604705
- Scarborough, Hollis S. 1990a. Index of productive syntax. – *Applied Psycholinguistics*, 11, 11–22. doi:10.1017/S0142716400008262
- Scarborough, Hollis S. 1990b. Very early syntactic deficits in dyslexic children. – *Child Development*, 61, 1728–1743. doi:10.2307/1130834

- Scarborough, Hollis S.; Rescorla, Leslie; Tager-Flusberg, Helen; Fowler, Anne E.; Sudhalter, Vicki 1991. The relation of utterance length to grammatical complexity in normal and language-disordered groups. – *Applied Psycholinguistics*, 12, 23–45. doi:10.1017/S014271640000936X
- Silverman, Stacey W.; Ratner, Nan Bernstein 1997. Syntactic complexity, fluency, and accuracy of sentence imitations in adolescents. – *Journal of Speech, Language and Hearing Research*, 40, 95–107.
- Smith, Carlota 1988. Factors of linguistic complexity and performance. – Alice Davison, Georgia M. Green (Eds.). *Linguistic Complexity and Text Comprehension: Readability Issues Reconsidered*. Hillsdale, NJ: Lawrence Erlbaum Associates, 247–279.
- Smith, Carlota; van Kleeck, Anne 1986. Linguistic complexity and performance. – *Journal of Child Language*, 13, 389–408. doi:10.1017/S0305000900008126
- Tomasello, Michael 2000. First steps toward a usage-based theory of language acquisition. – *Cognitive Linguistics*, 11, 61–82. doi:10.1515/cogl.2001.012
- Tomasello, Michael 2003. *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Yaruss, J. Scott 1999. Utterance length, syntactic complexity, and childhood stuttering. – *Journal of Speech, Language, and Hearing Research*, 42, 329–345.

Lea Nieminen (University of Jyväskylä) research interests include the structural complexity of language and the acquisition of morphosyntax. She has taught general linguistics and research methodology to students of Finnish Sign Language at Jyväskylä University and has also taught courses focusing on language acquisition at Helsinki and Jyväskylä Universities.
lea.s.m.nieminen@jyu.fi

VKP JA IPSyn ABSOLUUTSE KEERUKUSE MÕÕTJAINA

Lea Nieminen

Jyväskylä Ülikool

Artiklis vaadeldakse, kuidas suhtestuvad struktuurilise keerukuse mõõtmiseks kasutatavad vahendid, nagu VKP (väljendite keskmine pikkus, inglise keeles MLU) ja produktiivse süntaksi indeks IPSyn, millega mõõdeti 30 kuu vanuste soome laste kasutatud spontaanseid lauseid. Võrdluse eesmärk oli teha kindlaks, milliseid struktuurilise keerukuse aspekte kumbki mõõdik näitab. Struktuurilise keerukuse defineerimiseks on lähemalt peatunud nii absoluutsel kui ka mitmeta-sandilisel lähenemisel lause keerukusele, lisaks on neid vahendeid rakendatud ka lauseanalüüsis.

Mõõdikute kasutamise ja struktuurilise keerukuse muutumise tulemuste analüüs näitas, et VKP ja IPSyn on mõlemad struktuurilise keerukuse mõõtmiseks kasutatavad, kuid ainult juhul, kui neid rakendatakse samaaegselt. Mõõdikud on suunatud keelekasutuse eri aspektidele ning mõlema mõõdiku abil saadud tulemusi tuleks interpreteerida omavahelises suhtes.

Uurimuse tulemustel jõuti järeldusele, et lastel toimub morfoloogiline areng kõigepealt ühe süntaktilise komponendi piires, seejärel jätkub morfoloogiline areng ühes ja süntaktiline areng mõnes teises komponendis ja alles lõpuks areneb üks ja sama keeleline komponent nii morfoloogiliselt kui ka süntaktiliselt.

Seetõttu võib väita, et struktuuriline keerukus kasvab eri arengufaasidel eri keelelistel tasanditel. Kumbki siinses kirjutises esitatud mõõdik ei näita aga keelelise arengu kõiki aspekte, seetõttu ongi vajalik kasutada korraga mõlemat mõõdikut ning saadud tulemusi põhjalikult interpreteerida.

Võtmesõnad: morfosüntaks, keeleomandamine, struktuuriline keerukus, lastekeel, soome keel