Figurative Drawing Abilities Map onto Distinct Cognitive Mechanism from Non-Figurative Abilities in 77,000 Participants with Neurodevelopmental Disorders

Andrey Vyshedskiy[1*], Rohan Venkatesh[2], Edward Khokhlovich[3]

1. Boston University, Metropolitan College, Boston, Massachusetts, USA

2. University of Missouri-Kansas City, Kansas City, MO, USA

3. Independent Researcher, USA

*Corresponding author: Andrey Vyshedskiy, Ph.D., Boston University, Boston, USA, Tel: +1 (617) 433-7724; E-mail: vysha@bu.edu

Abstract

Non-figurative motifs (parallel lines, hand stencils) constitute the earliest known human art, dating to ~400,000 years ago. In contrast, figurative art—depicting recognizable animals and humans—emerges abruptly ~45,000 ya. We hypothesized that this temporal dissociation reflects the sequential evolution of distinct neurocognitive mechanisms underlying language and visual manipulation. Building on our prior taxonomic framework that identified distinct expressive and receptive language mechanisms, we examined their co-occurrence with drawing abilities in individuals with neurodevelopmental disorders. Results revealed specific dissociations: figurative drawing emerged in parallel with the Syntactic Mechanism (responsible for syntactic comprehension). In contrast, non-figurative art correlated with the Modifier Mechanism (mediating the integration of nouns with adjectives). These dissociations suggest that the cognitive architecture required for figurative depiction is tightly linked to the same prefrontal-parieto-temporal system that supports the Syntactic Mechanism. Conversely, earlier-emerging non-figurative artistic traditions appear to rely primarily on a more basic integrative capacity for the Modifier Mechanism. We therefore propose that the appearance of non-figurative art in the Middle Pleistocene likely coincided with the evolutionary emergence of the Modifier Mechanism, whereas the appearance of figurative art reflects the later acquisition of the Syntactic Mechanism and its associated capacity for comprehension of syntactic, relational sentence structure.

Keywords: symbolic thinking, figurative art, figurative art, symbolic drawing, figurative drawing, human evolution, language evolution

Introduction

Drawing has long been used as a window into symbolic abilities, both for monitoring child development and for diagnosing dementia [1–3]. The study of human evolution is no exception, with

symbolic capacities frequently inferred from evidence of drawing. Discoveries of artistic artifacts in Eurasia that predate the migration of *Homo sapiens* have sparked significant excitement in the scientific community. Pre–*Homo sapiens* artistic artifacts include use of pigments [4], perforated shells [4–8], line marks on stones and shells [9,10], geometrical figures [11], and hand stencils painted on cave walls [12–18]. While these artifacts may be symbolic, they are non-figurative—they do not represent clear images of humans, animals, or other recognizable objects.

Figurative art appears in Eurasia only after the migration of *Homo sapience* out-of-Africa, around 70,000 years ago. The earliest examples of figurative art include a hunting scene depicting part-human, part-animal figures from the limestone cave of Leang Bulu' Sipong 4 (Sulawesi, Indonesia), dated to 44,000 years ago [19]; a painting of an animal at a limestone cave in Indonesian Borneo, dated to 40,000 years ago [20]; a drawing of a "deer-pig" from the Indonesian island of Sulawesi, dated to 40,000 years ago [21]; an engraving of an extinct wild cow (aurochs) from the Dordogne region of France, dated 38,000 years ago [22]; and the lion-man sculpture excavated from the caves of Lone valley in Germany, dated to 39,000 years ago [23].

Does the transition from non-figurative to figurative traditions reflect a fundamental shift in symbolic cognition? One approach to examining the relationship between artistic capacities and cognition is the study of contemporary individuals with neurodevelopmental disorders. Many such individuals harbor deleterious mutations in genes implicated in cognitive and language evolution, including FOXP2 [24–26] or SRGAP2A [27], and may thus model features of ancestral neurocognitive architectures.

The motivation for this study was two-fold. First, we had access to the largest database of over 100,000 individuals exhibiting a range of symbolic abilities, with parent-reported figurative and non-figurative drawing skills. Second, previous work has demonstrated distinct clustering of receptive and expressive language abilities [28–31]. Access to this data enabled the present investigation of associations between drawing skills and language phenotypes. Specifically, we examine how drawing abilities co-occur with different language phenotypes. If drawing abilities were independent of language phenotypes—like unrelated traits such as hyperactivity— they would be expected to occur with equal frequency across language levels and form a cluster separate from the receptive and expressive language clusters. Conversely, a strong association between a drawing ability and a specific language phenotype would indicate a close maturational link between symbolic drawing skills and language proficiency.

Methods

Study Participants

Participants were children and adolescents using a language therapy app that was made freely available at all major app stores in September 2015 [32–36]. The app provides various structured language comprehension therapy exercises and is primarily used by caregivers of children with language impairments. Most of the caregivers are presumed to be parents. Once the app was downloaded, caregivers were asked to register and to provide demographic details, including the child's diagnosis and age. Caregivers completed a 133-item questionnaire (77-item Autism Treatment Evaluation Checklist (ATEC) [37] (Supplementary Tables 1-4); 20-item Mental Synthesis

Evaluation Checklist (MSEC) [38] (Supplementary Table 5); 10-item screen time checklist [39]; 25-item diet checklist [40]; and 1-item parent education survey) approximately every three months.

All 26 language-related items (15 receptive, Table 1, and 11 expressive, Table 2) available in the 133-item questionnaire were included in the cluster analysis as in previously published articles [28–31]. Response options were: very true (0 points), somewhat true (1 point), and not true (2 points).

The inclusion criteria for this study remained consistent with those of previous studies [28–31]: absence of seizures (which commonly result in intermittent, unstable language deficits [41]), absence of serious and moderate sleep problems (which are also associated with intermittent, unstable language deficits [42]), age range of 4 to 22 years (the lower age cutoff was chosen to ensure that participants were exposed to complete set of items listed in Table 1 [43], whereas the upper age threshold was constrained by the scarcity of older participants).

Table 3 reports participants' diagnoses as communicated by caregivers. Autism level (mild/Level 1, moderate/Level 2, or severe/Level 3) was reported by caregivers. Pervasive Developmental Disorder and Asperger Syndrome were combined with mild autism for analysis as recommended by DSM-5 [44]. A good reliability of such parent-reported diagnosis has been previously demonstrated [45].

When caregivers completed several evaluations, the last evaluation was used for analysis. Thus, the study included a total of 77,616 participants, the average age was 6.6±2.7 years (range of 4 to 22 years), 76.3% participants were males (Table 3). The education level of participants' parents was the following: 90.8% with at least a high school diploma, 68.7% with at least college education, 35.9% with at least a master's, and 5.7% with a doctorate.

The study was conducted in accordance with the Declaration of Helsinki [46]. Informed consent was obtained from the caregivers of all participants. The study protocol was approved by the Biomedical Research Alliance of New York (BRANY) LLC Institutional Review Board.

Drawing assessments

The evaluation included three questions about participants' drawing abilities. One item assessed non-figurative drawing: "[My child] does drawing, coloring, art;" and two items assessed figurative drawing ability: "[My child] draws a VARIETY of RECOGNIZABLE images (objects, people, animals, etc.);" and "[My child] can draw a NOVEL image following YOUR description (e.g. a three-headed horse)." Response options were: "very true," "somewhat true," and "not true."

Statistics

Unsupervised Hierarchical Cluster Analysis (UHCA) is a robust method for detecting patterns of co-occurring abilities within noisy data. It is fully data-driven, requiring no a priori assumptions about the number or structure of clusters, and thus enables the identification of naturally emerging groupings based solely on the statistical relationships among variables. UHCA was performed using Ward's agglomeration method with a Euclidean distance metric.

An additional statistical approach, Principal Component Analysis (PCA), was used to confirm the UHCA findings. PCA employs a fundamentally different mathematical framework for dimensionality reduction, transforming high-dimensional, noisy measures into a compact set of orthogonal components that capture the dominant sources of variance in the data. This data-driven projection spatially organizes abilities that co-occur more frequently closer together in a two-dimensional component space, providing an independent convergence test for the clustering structure.

Analysis was performed in R, freely available language for statistical computing [47].

Code and data can be downloaded from https://doi.org/10.17605/OSF.IO/2QK5B???.

Results

Clustering analysis of language comprehension abilities

Fifteen language comprehension abilities were assessed by caregivers (Table 1). To examine patterns of co-occurrence among these abilities, we applied unsupervised hierarchical cluster analysis (UHCA)—a data-driven method that produces tree-like diagrams, called dendrograms, which visually represent the hierarchical relationships between clusters of items. Abilities that frequently co-occur are automatically positioned closer together, while those that co-occur less often, appear farther apart.

The UHCA-derived dendrogram (Figure 1A) replicated the hierarchical cluster structure reported in our prior work [28–30]. The first cluster included "*[My child] knows his/her name*", "*[My child] responds to 'No' or 'Stop',*" *"[My child] responds to praise,"* and "*[My child] follows some commands*" (items 1 to 4 in Table 1) and was previously termed the Command Mechanism.

The second cluster included "*[My child] understands color and size modifiers,"* "*[My child] understands several modifiers in a sentence,"* "*[My child] understands size superlatives,"* and "*[My child] understands numbers"* (items 5 to 8 in Table 1) and was designated as the Modifier Mechanism in our prior work.

The third cluster included "*[My child] understands spatial prepositions,"* "*[My child] understands verb tenses,"* "*[My child] understands flexible syntax,"* "*[My child] understands possessive pronouns,"* "*[My child] understands explanations about people and situations,"* "*[My child] understands simple stories,"* and "*[My child] understands elaborate fairytales"* (items 9 to 15 in Table 1) and was previously termed the Syntactic Mechanism.

Principal component analysis (PCA) (Figure 1B) also showed a clear separation between the Command, Modifier, and Syntactic clusters. Each cluster occupied a unique region in the principal component space, with no overlap observed. This separation suggests that the underlying structure of the data effectively captures the functional differences among the clusters, providing strong evidence for their distinct classification.

Our previous analysis using both UHCA and PCA demonstrated that the Command, Modifier, and Syntactic clusters were highly stable across multiple evaluation methods (Ward, Average, Complete, McQuitty) [28,29], age groups (4 to 6 years, 6 to 12 years, 12 to 22 years) [28,29], time

points (first vs. last evaluation) [28,29], genders [28], spoken-languages (English, Spanish, Portuguese, Italian, Russian, Chinese, French, German, or Korean) [30], diagnostic categories [28,30], and levels of parental education [28].

The reasons for the emergence of three distinct clusters of co-occurring abilities become clear when viewed through the lens of developmental dynamics. Some individuals with language impairments plateau at the Command level, others at the Modifier level, while only a subset progress to the Syntactic level [48]. Individuals who plateau at the Command level typically acquire the full set of Command abilities; those who plateau at the Modifier level acquire both Command and Modifier abilities; and only individuals who reach the Syntactic level go on to acquire syntactic abilities. This developmental stratification explains the observed clustering pattern: the four Command abilities (items 1–4 in Table 1) reliably co-occur and are grouped by UHCA and PCA into the Command cluster; the four Modifier skills (items 5–8) co-occur and form the Modifier cluster; and the seven Syntactic abilities (items 9–15) co-occur and constitute the Syntactic cluster.

As a control we calculated UHCA and PCA of the 15 language comprehension abilities along with the *hyperactivity* (Supplementary Figure 1) items. Since this item is not related to language, it would be expected to form a separate cluster. As anticipated, both UHCA and PCA grouped *hyperactivity* into its own cluster at a significant distance from the three comprehension clusters, indicating random co-occurrence with the language abilities, and, thus, validating the effectiveness of both clustering techniques.

Clustering analysis of expressive language abilities

Eleven expressive language abilities were assessed by caregivers (Table 2). The dendrogram generated by the UHCA of these abilities (Figure 2A) replicated the hierarchical cluster structure reported in our prior work [31]. Three distinct clusters emerged. The first cluster, previously termed Single-Word, comprised the abilities "*[My child] can use one word at a time*" and "*[My child] knows 10 or more words*" (Table 2, items 1 and 2).

The second cluster, termed Single-Sentence, included the abilities "*[My child] can use 2 words at a time*", "*[My child] can use three words at a time*", "*[My child] can use sentences with 4 or more words*", "*[My child] explains what they want*", and "*[My child] speech tends to be meaningful/relevant*" (items 3 to 7).

The third cluster, termed the Multi-Sentence, encompasses the abilities "*[My child] carries on fairly good conversation*", "*[My child] has normal ability to communicate for their age*", "*[My child] asks meaningful questions*", and "*[My child] often uses several successive sentences"* (items 8 to 11). PCA (Figure 2B) also showed a clear separation between the Command, Modifier, and Syntactic clusters. Previous analysis explored the stability of the expressive language clusters [31].

The reasons for the emergence of the three expressive language clusters mirror those underlying the three comprehension clusters. Specifically, some children plateau at the Single-Word level, others at the Single-Sentence level, while the majority progress to the Multi-Sentence level (manuscript in preparations). Consequently, clustering analysis of large cohorts of children with

language impairments consistently identify these developmental plateaus as sets of co-occurring abilities, reflecting the natural grouping of linguistic competencies.

As a control we calculated UHCA and PCA of the 11 expressive language abilities along with the *hyperactivity* (Supplementary Figure 2) item. This item is not related to speech and was therefore expected to form a separate cluster. As anticipated, both UHCA and PCA grouped *hyperactivity* into its own cluster at a significant distance from the three expressive clusters, validating the effectiveness of both clustering techniques.

Clustering analysis of language comprehension abilities together with expressive language abilities

Figure 3A depicts the UHCA dendrogram of the 15 language comprehension items (Table 1) together with the 11 expressive language items (Table 2). As in the previous study [31], both the UHCA (Figure 3A) and the PCA (Figure 3B) show six clusters of items. These clusters are identical to those detected by separate clustering of receptive (Figure 1) and expressive (Figure 2) items: three language comprehension clusters (Command, Modifier, and Syntactic) and three expressive language clusters (Single-Word, Single-Sentence, and Multi-Sentence).

As a control we calculated UHCA and PCA of the 15 language comprehension abilities together with the 11 expressive language items along with the *hyperactivity* item, which is not expected to be related to any particular language ability and therefore should cluster into its own group. As expected, both UHCA and PCA clustered the hyperactivity item into its own group at a significant distance from the six language clusters, validating the effectiveness of both clustering techniques (Figure S3).

Clustering analysis of receptive-expressive language abilities together with drawing abilities

Next, we calculated UHCA and PCA of the 15 receptive and 11 expressive abilities along with the non-figurative drawing item "[My child] does drawing, coloring, art" (Figure 4). Both UHCA and PCA analysis clustered non-figurative drawing with the Modifier Mechanism indicating their common co-occurrence.

Conversely, both UHCA and PCA clustered the figurative-drawing item "[My child] draws a VARIETY of RECOGNIZABLE images (objects, people, animals, etc.)" with the Syntactic Mechanism (Figure 5).

Finally, UHCA and PCA grouped the second figurative-drawing item "[My child] can draw a NOVEL image following YOUR description (e.g. a three-headed horse)" also with the Syntactic Mechanism (Figure 6).

Interpreted in a developmental context, these findings suggest that individuals with language impairments who plateau at the Modifier level typically show a propensity for non-figurative drawing but not for figurative drawing. In contrast, individuals who progress to the Syntactic level usually acquire the ability to draw figuratively as well.

Stability of drawing abilities associations with language clusters

We examined the stability of the association between drawing abilities and language clusters across different age groups (4-6 years, n=41,188; 6-12 years, n=32,074; and 12-22 years, n=4,354), levels of autism severity (severe, n=11,428; moderate, n=19,844; and mild, n=31,072), native language (English, n=26,020 and non-English, n=51,596), and participants' sex (female, n=18,413 and male, n=57,398) (Supplementary Figures 4-31).

Figurative drawing ability, as reported by caregivers in response to the items "[My child] draws a VARIETY of RECOGNIZABLE images (objects, people, animals, etc.)" and "[My child] can draw a NOVEL image following YOUR description (e.g. a three-headed horse)," consistently clustered with the Syntactic Mechanism (Supplementary Figures 14-31).

Conversely, the non-figurative drawing ability, as reported by caregivers in response to the item "[My child] does drawing, coloring, art," never clustered with the Syntactic Mechanism or any of the three expressive mechanisms. PCA consistently positioned this item closer to the Modifier Phenotype (Supplementary Figures 5-13B) and UHCA clustered it with the Modifier Mechanism in most strata (Supplementary Figures 5, 6, 7, 8, 10, 11, and 13A). These results indicate that non-figurative drawing typically emerges in individuals with the Modifier Phenotype. In three subpopulations, UHCA clustered this item with the Command Mechanism (Supplementary Figures 4, 9, 12A), possibly reflecting insufficient resolution of UHCA to detect clustering in smaller subgroups.

Discussion

When biological abnormalities slow-down or prevent full language attainment, receptive and expressive language development follows a distinct step-wise pattern. Receptive comprehension progresses through the Command → Modifier → Syntactic Mechanisms, while expressive language advances through Nonverbal → Single-Word → Single-Sentence → Multi-Sentence production. When development stalls, it typically does so at a full step rather than at an intermediate or half-step [48]. The results of this study indicate that drawing abilities are strongly associated with specific levels of language comprehension, rather than a stage in expressive language development. The presence of the Syntactic Mechanism is closely linked to figurative drawing: children who do not acquire the Syntactic Mechanism also typically fail to acquire figurative drawing ability. Likewise, the presence of the Modifier Mechanism is strongly associated with non-figurative drawing, and children who do not attain the Modifier Mechanism (i.e. retain the Command Phenotype) also usually do not acquire even non-figurative drawing skills.

These results are consistent with data from neurotypical children. Neurotypical children typically acquire the Command Mechanism around age 2, the Modifier Mechanism around age 3, and the Syntactic Mechanism around age 4 years [49]. Their drawing abilities follow a parallel developmental patterns, with non-figurative scribbling emerging around age 3 and figurative "pre-schematic" drawings appearing by about age 4 [50]. However, the associations between comprehension and drawing phenotypes are difficult to study in neurotypical populations because children pass through all language levels rapidly and are often introduced to drawing after they acquire the Syntactic Phenotype. In contrast, children with language impairments often

remain at inferior levels for years [48], providing an ideal population in which to examine the relationship between language mechanisms and drawing abilities.

Implications for the evolution of human language

Syntactic comprehension does not fossilize, making its evolutionary acquisition difficult to determine. Traditionally, it has been examined through the lens of symbolic thinking, which includes drawing ability. This study's findings suggest that not all drawing abilities depend equally on the syntactic level: figurative drawing is strongly associated with the Syntactic Phenotype, whereas non-figurative drawing is primarily associated with the Modifier Phenotype.

This pattern may help explain the clear temporal sequence observed in the archaeological record. Non-figurative motifs—such as parallel lines, dots, and hand stencils—represent the earliest known forms of human visual art, appearing as early as ~400,000 years ago. In contrast, figurative art, depicting recognizable animals, humans, or scenes, emerges abruptly around 45,000 years ago. Our results suggest that this striking temporal dissociation may reflect the sequential evolutionary emergence of distinct neurocognitive mechanisms underlying the Modifier and the Syntactic Phenotypes.

Limitations

The main limitation of this study is its associational nature: a syntactic comprehension phenotype does not directly cause drawing abilities. Some individuals may possess the Syntactic Comprehension Phenotype but fail to exhibit it in figurative drawing due to limited cultural encouragement or fine-motor constraints. Conversely, others may lack the Syntactic Phenotype yet display some figurative drawing ability. Autistic savants are particularly notable in this regard, as their exceptional figurative drawing abilities often contrast sharply with markedly limited syntactic skills [51,52]. Importantly, the mechanisms supporting figurative drawing in autistic savants may differ from the mental combinatorial pre-planning typically used by neurotypical individuals. For instance, the contours of certain animals can become automatized through extensive practice. In such cases, drawing an animal resembles 'writing a signature'—executed with minimal conscious planning, enabling drawing of the full outline of the animal in a single fluid motion. In our parent-report assessments, we attempted to mitigate the potential influence of automated drawing by emphasizing 'VARIETY of animals' in the item, "My child draws a VARIETY of RECOGNIZABLE images (objects, people, animals, etc.)." However, we cannot determine how many participants in our cohort may have reported the presence of an automatized drawing skill.

Funding

This research received no external funding.

Acknowledgements

We wish to thank all participants' caregivers who found time to complete children's assessments. The language therapy app used to collect the data presented in this manuscript was made possible by the contributions of Rita Dunn, Alexander Faisman, Jonah Elgart, Lisa Lokshina, and Yulia Dumov.

Author contributions

AV and EK designed the study. RV, EK, and AV analyzed the data. EM and AV wrote the paper.

Competing Interests

Authors declare no competing interests.

References

1. Beyn, E. S. & Knyazeva, G. R. The problem of prosopagnosia. *J. Neurol. Neurosurg. Psychiatry* **25**, 154 (1962).
2. Levine, D. N., Warach, J. & Farah, M. Two visual systems in mental imagery: Dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions. *Neurology* **35**, 1010–1010 (1985).
3. Thomas, G. V. & Silk, A. M. *An Introduction to the Psychology of Children's Drawings.* (New York University Press, 1990).
4. Zilhão, J. *et al.* Symbolic use of marine shells and mineral pigments by Iberian Neandertals. *Proc. Natl. Acad. Sci.* **107**, 1023–1028 (2010).
5. Bouzouggar, A. *et al.* 82,000-year-old shell beads from North Africa and implications for the origins of modern human behavior. *Proc. Natl. Acad. Sci.* **104**, 9964–9969 (2007).
6. Henshilwood, C., d'Errico, F., Vanhaeren, M., Van Niekerk, K. & Jacobs, Z. Middle stone age shell beads from South Africa. *Science* **304**, 404–404 (2004).
7. d'Errico, F., Henshilwood, C., Vanhaeren, M. & Van Niekerk, K. Nassarius kraussianus shell beads from Blombos Cave: evidence for symbolic behaviour in the Middle Stone Age. *J. Hum. Evol.* **48**, 3–24 (2005).

8. Sehasseh, E. M. *et al.* Early Middle Stone Age personal ornaments from Bizmoune Cave, Essaouira, Morocco. *Sci. Adv.* **7**, eabi8620 (2021).
9. Henshilwood, C. S. *et al.* An abstract drawing from the 73,000-year-old levels at Blombos Cave, South Africa. *Nature* **562**, 115–118 (2018).
10. Joordens, J. C. *et al.* Homo erectus at Trinil on Java used shells for tool production and engraving. *Nature* **518**, 228–231 (2015).
11. Marquet, J.-C. *et al.* The earliest unambiguous neanderthal engravings on cave walls: La Roche-Cotard, Loire valley, France. *PLoS One* **18**, e0286568 (2023).
12. Henshilwood, C. S., d'Errico, F. & Watts, I. Engraved ochres from the middle stone age levels at Blombos Cave, South Africa. *J. Hum. Evol.* **57**, 27–47 (2009).
13. Hoffmann, D. L. *et al.* U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science* **359**, 912–915 (2018).
14. Standish, C. D. *et al.* The age of hand stencils in Maltravieso cave (Extremadura, Spain) established by U-Th dating, and its implications for the early development of art. *J. Archaeol. Sci. Rep.* **61**, 104891 (2025).
15. Pettitt, P. *et al.* Are hand stencils in European cave art older than we think? An evaluation of the existing data and their potential implications. *Prehist. Art Prehist. Cult. Stud. Honour Profr. Rodrigo Balbín-Behrmann* 31–43 (2015).
16. Navarro, V. F., Casares, D. F., Martínez, D. G. & Maidagan, D. G. Decoding Palaeolithic Hand Stencils: Age and Sex Identification Through Geometric Morphometrics. *J. Archaeol. Method Theory* **32**, 24 (2025).

17. Ripoll, S. *et al.* Hands Stencils in El Castillo Cave (Puente Viesgo, Cantabria, Spain). An Interdisciplinary Study. in *Proceedings of the prehistoric society* vol. 87 51–71 (Cambridge University Press, 2021).
18. Fernández-Navarro, V., Spaey, O. & Garate, D. Reevaluating Hand Stencil Phenomena in Cave Art: A Step Forward towards the Characterization of Symbolic Patterns during the Upper Palaeolithic in Europe. *Camb. Archaeol. J.* **1**, 21 (2026).
19. Aubert, M. *et al.* Earliest hunting scene in prehistoric art. *Nature* **576**, 442–445 (2019).
20. Aubert, M. *et al.* Palaeolithic cave art in Borneo. *Nature* **564**, 254–257 (2018).
21. Aubert, M. *et al.* Pleistocene cave art from Sulawesi, Indonesia. *Nature* **514**, 223–227 (2014).
22. Bourrillon, R. *et al.* A new Aurignacian engraving from Abri Blanchard, France: Implications for understanding Aurignacian graphic expression in Western and Central Europe. *Quat. Int.* http://www.sciencedirect.com/science/article/pii/S1040618216305766 (2017).
23. Dalton, R. *Lion Man Takes Pride of Place as Oldest Statue*. (Nature Publishing Group, 2003).
24. Moreno Campos, V. & Benítez-Burraco, A. Communication deficits in a case of a deletion in 7q31.1-q31.33 encompassing *FOXP2*. *Clin. Linguist. Phon.* **37**, 1157–1170 (2023).
25. Morison, L. D., Braden, R., Amor, D. J. & Morgan, A. T. Speech and Language Disorders Associated With 7q31 Deletions Implicating *FOXP2* . *Am. J. Med. Genet. A.* e64190 (2025) doi:10.1002/ajmg.a.64190.
26. Kuo, H.-Y. *et al.* Speech-and language-linked FOXP2 mutation targets protein motors in striatal neurons. *Brain* **146**, 3542–3557 (2023).

27. Martins, P. T., Marí, M. & Boeckx, C. SRGAP2 and the gradual evolution of the modern human language faculty. *J. Lang. Evol.* **3**, 67–78 (2018).
28. Vyshedskiy, A., Venkatesh, R., Khokhlovich, E. & Satik, D. Three mechanisms of language comprehension are revealed through cluster analysis of individuals with language deficits. *Npj Sci. Learn.* **9**, 1–12 (2024).
29. Vyshedskiy, A., Venkatesh, R. & Khokhlovich, E. Are there distinct levels of language comprehension in autistic individuals – cluster analysis. *Npj Ment. Health Res.* **3**, (2024).
30. Murphy, E., Venkatesh, R., Khokhlovich, E. & Vyshedskiy, A. Merge-based syntax is mediated by distinct neurocognitive mechanisms in 84,000 individuals with language deficits across nine languages. *Sci. Rep.* **15**, 44444 (2025).
31. Gankin, K., Venkatesh, R., Khokhlovich, E. & Vyshedskiy, A. Receptive and expressive language phenotyping in over 62,000 autistic Individuals. https://www.researchsquare.com/article/rs-8466583/latest (2026).
32. Vyshedskiy, A. & Dunn, R. Mental Imagery Therapy for Autism (MITA)-An Early Intervention Computerized Brain Training Program for Children with ASD. *Autism Open Access* **5**, 2 (2015).
33. Dunn, R. *et al.* Comparison of performance on verbal and nonverbal multiple-cue responding tasks in children with ASD. *Autism Open Access* **7**, 218 (2017).
34. Dunn, R. *et al.* Tablet-Based Cognitive Exercises as an Early Parent-Administered Intervention Tool for Toddlers with Autism - Evidence from a Field Study. *Clin. Psychiatry* **3**, (2017).

35. Dunn, R. *et al.* Children With Autism Appear To Benefit From Parent-Administered Computerized Cognitive And Language Exercises Independent Of the Child's Age Or Autism Severity. *Autism Open Access* **7**, (2017).
36. Vyshedskiy, A. *et al.* Novel prefrontal synthesis intervention improves language in children with autism. *Healthcare* **8**, 566 (2020).
37. Rimland, B. & Edelson, S. M. Autism treatment evaluation checklist (ATEC). *Autism Res. Inst. San Diego CA* http://www.autism.com (1999).
38. Braverman, J., Dunn, R. & Vyshedskiy, A. Development of the Mental Synthesis Evaluation Checklist (MSEC): A Parent-Report Tool for Mental Synthesis Ability Assessment in Children with Language Delay. *Children* **5**, 62 (2018).
39. Fridberg, E., Khokhlovich, E. & Vyshedskiy, A. Watching Videos and Television Is Related to a Lower Development of Complex Language Comprehension in Young Children with Autism. in *Healthcare* vol. 9 423 (Multidisciplinary Digital Publishing Institute, 2021).
40. Acosta, A., Khokhlovich, E., Reis, H. & Vyshedskiy, A. Dietary factors impact developmental trajectories in young autistic children. *J. Autism Dev. Disord.* https://doi.org/10.1007/s10803-023-06074-8 (2023) doi:10.1007/s10803-023-06074-8.
41. Forman, P., Khokhlovich, E. & Vyshedskiy, A. Longitudinal Developmental Trajectories in Young Autistic Children Presenting with Seizures, Compared to those Presenting without Seizures, Gathered via Parent-report Using a Mobile Application. *J. Dev. Phys. Disabil.* https://doi.org/10.1007/s10882-022-09851-y (2022) doi:10.1007/s10882-022-09851-y.
42. Levin, J., Khokhlovich, E. & Vyshedskiy, A. Longitudinal developmental trajectories in young autistic children presenting with sleep problems, compared to those presenting without

sleep problems, gathered via parent-report using a mobile application. *Res. Autism Spectr. Disord.* **97**, 102024 (2022).

43. Arnold, M. & Vyshedskiy, A. Combinatorial language parent-report score differs significantly between typically developing children and those with Autism Spectrum Disorders. *J. Autism Dev. Disord.* https://doi.org//10.1007/s10803-022-05769-8 (2022) doi:/10.1007/s10803-022-05769-8.
44. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
45. Jagadeesan, P., Kabbani, A. & Vyshedskiy, A. Parent-reported assessment scores reflect ASD severity level in 2- to 7- year-old children. *Children* **9**, 701 (2022).
46. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* **310**, 2191–2194 (2013).
47. R Foundation for Statistical Computing. R: A language and environment for statistical computing. (2021).
48. Venkatesh, R., Nowakowski, A., Khokhlovich, E. & Vyshedskiy, A. Longitudinal trajectories across the Command, Modifier, and Syntactic Phenotypes of language comprehension in over 6,000 autistic children. Preprint at https://doi.org/10.64898/2025.12.19.25342690 (2025).
49. Vyshedskiy, A. *et al.* Language Comprehension Developmental Milestones in Typically Developing Children Assessed by the New Language Phenotype Assessment (LPA). *Child. Basel Switz.* **12**, 793 (2025).

50. Piaget, J. *Play, Dreams and Imitation in Childhood*. (W. W. Norton & Co. Orig.: La formation du symbole chez l'enfant. Neuchatel: Delachaux & Niestlé., 1945).

51. Pring, L., Ryder, N., Crane, L. & Hermelin, B. Creativity in savant artists with autism. *Autism* **16**, 45–57 (2012).

52. Sacks, O. *The Man Who Mistook His Wife for a Hat: A Revealing Exploration of the Mysteries of the Human Mind*. vol. 19 (Pan Macmillan, 2014).

		Language comprehension items (verbatim)	Abbreviations used in dendrograms
Command Mechanism	1	Knows own name	Knows Name
	2	Responds to 'No' or 'Stop'	No and Stop
	3	Responds to praise	Resp. to Praise
	4	Can follow some commands	Commands
Modifier Mechanism	5	Understands some simple modifiers (i.e., green apple vs. red apple or big apple vs. small apple)	Color or Size / Modifiers
	6	Understands several modifiers in a sentence (i.e., small green apple)	Two Modifiers
	7	Understands size (can select the largest/smallest object out of a collection of objects)	Size Superlatives
	8	Understands NUMBERS (i.e., two apples vs. three apples)	Numbers
Syntactic Mechanism	9	Understands spatial prepositions (i.e., put the apple ON TOP of the box vs. INSIDE the box vs. BEHIND the box)	Sp. Prepositions
	10	Understands verb tenses (i.e., I will eat an apple vs. I ate an apple)	Verb Tenses
	11	Understands simple stories that are read aloud	Simple Stories
	12	Understands elaborate fairytales that are read aloud (i.e., stories describing FANTASY creatures)	Elab. Fairytales
	13	Understands possessive pronouns (i.e., your apple vs. her apple)	Poss. Pronouns
	14	Understands the change in meaning when the order of words is changed (i.e., understands the difference between 'a cat ate a mouse' vs. 'a mouse ate a cat')	Flexible Syntax
	15	Understands explanations about people, objects or situations beyond the immediate surroundings (e.g., "Mom is walking the dog," "The snow has turned to water").	Explanations

Table 1. Three language comprehension mechanisms—Syntactic, Modifier, and Command—have been identified in previous studies [28–30]. When one mechanism is acquired, the entire range of associated comprehension abilities is also gained. The Command-level abilities (Items 1 to 4) are acquired first. The Modifier-level abilities (Items 5 to 8) are attained next. The Syntactic-level abilities (Items 9 to 15) are acquired last. The language comprehension items are presented exactly as surveyed with parents in both this and earlier studies. Response options were: very true, somewhat true, and not true. Items 1 to 3 were assessed as part of the Expressive Language ATEC [37] subscale 1; the rest of items were a part of the MSEC subscale [38].

		Expressive language items (verbatim)	Abbreviations used in dendrograms
Single-Word	1	[My child] Can use one word at a time (No!, Eat, Water, etc.)	Uses 1 word
	2	Knows 10 or more words	Knows 10+ words
Single-Sentence	2	Can use 2 words at a time (Don't want, Go home)	Uses 2 words
	3	Can use 3 words at a time (Want more milk)	Uses 3 words
	5	Can use sentences with 4 or more words	Uses 4+ words
	6	Explains what he/she wants	Explains what wants
	7	Speech tends to be meaningful/relevant	Speech meaningful
Multi-Sentence	2	Carries on fairly good conversation	Good conversation
	9	Has normal ability to communicate for his/her age	Normal communic.
	10	Asks meaningful questions	Asks questions
	11	Often uses several successive sentences	Uses several sentences

Table 2. Eleven expressive language items as they were posed to parents. Response options were: very true (0 points), somewhat true (1 point), and not true (2 points). All items were a part of the expressive language ATEC subscale 1. Single-Word, Single-Sentence, and Multi-Sentence clusters have been identified in previous work [31].

	Number of Participants	Percent of Total	Age, Mean(SD)	Percent Males
Mild Autism Spectrum Disorder (ASD)	31072	40.0	6.3(2.4)	76.8
Moderate ASD	19844	25.6	6.9(2.8)	80
Severe ASD	11498	14.8	7.6(3.3)	80.6
Specific Language Impairment	4928	6.3	6.0(2.2)	69.1
Mild Language Delay	4024	5.2	5.5(1.7)	67.4
ADHD	1649	2.1	6.4(2.2)	72.8
Down Syndrome	1450	1.9	8.5(3.5)	60.4
Other Genetic Disorder	1120	1.4	8.0(3.5)	57.7
Social Communication Disorder	805	1.0	6.3(2.3)	69.1
Sensory Processing Disorder	754	1.0	6.8(2.7)	70.2
Apraxia	472	0.6	6.8(2.9)	64.6
Total	77,616	100.0	6.6(2.7)	76.3

Table 3. Participants' diagnoses.

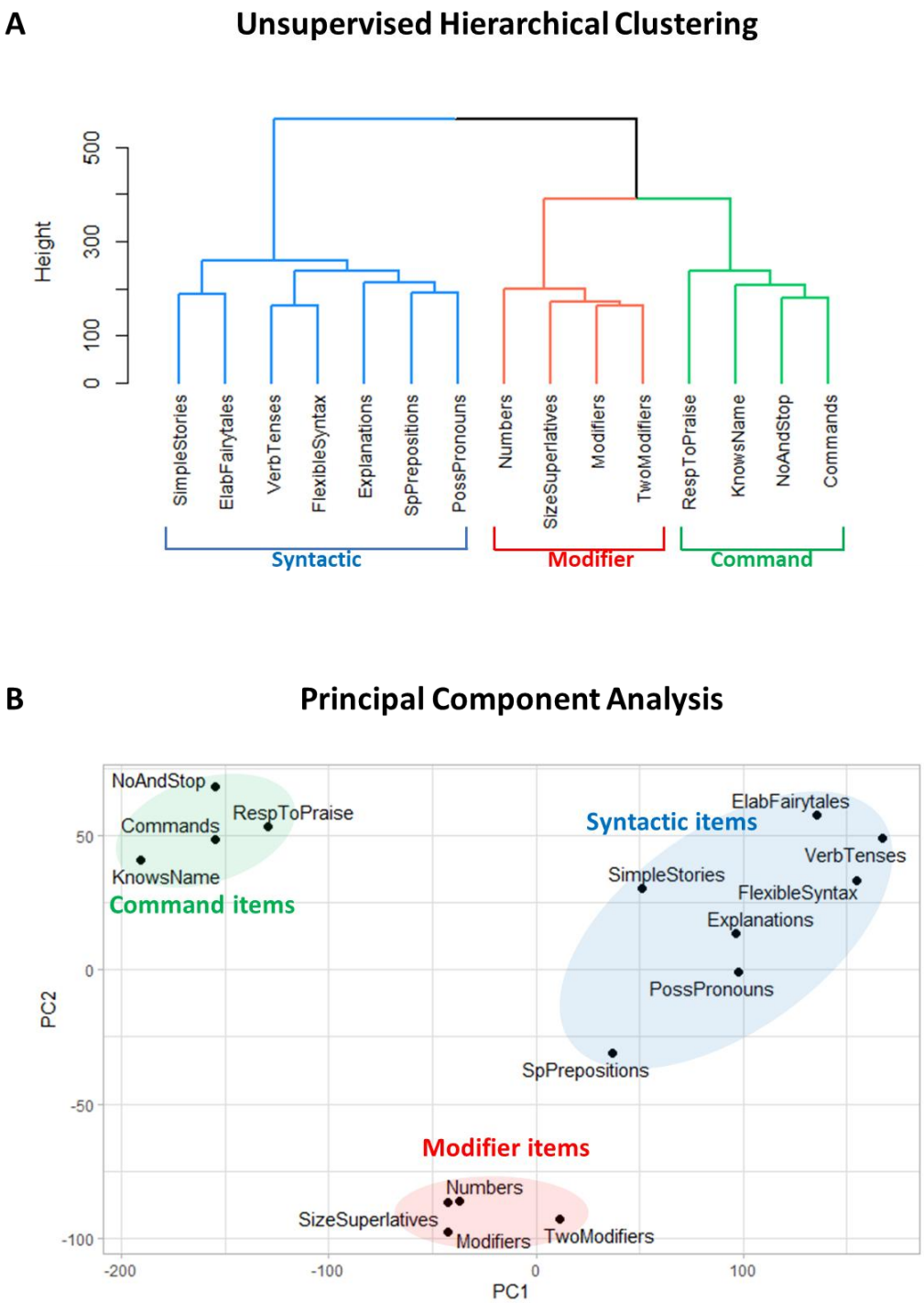

Figure 1. Clustering analysis of 15 language comprehension items. (A) The dendrogram representing the hierarchical clustering of language comprehension abilities. (B) Principal component analysis of language comprehension abilities shows clear separation between Command, Modifier, and Syntactic items. Principal component 1 accounts for 46.1% of the variance in the data. Principal component 2 accounts for 11.2% of the variance in the data.

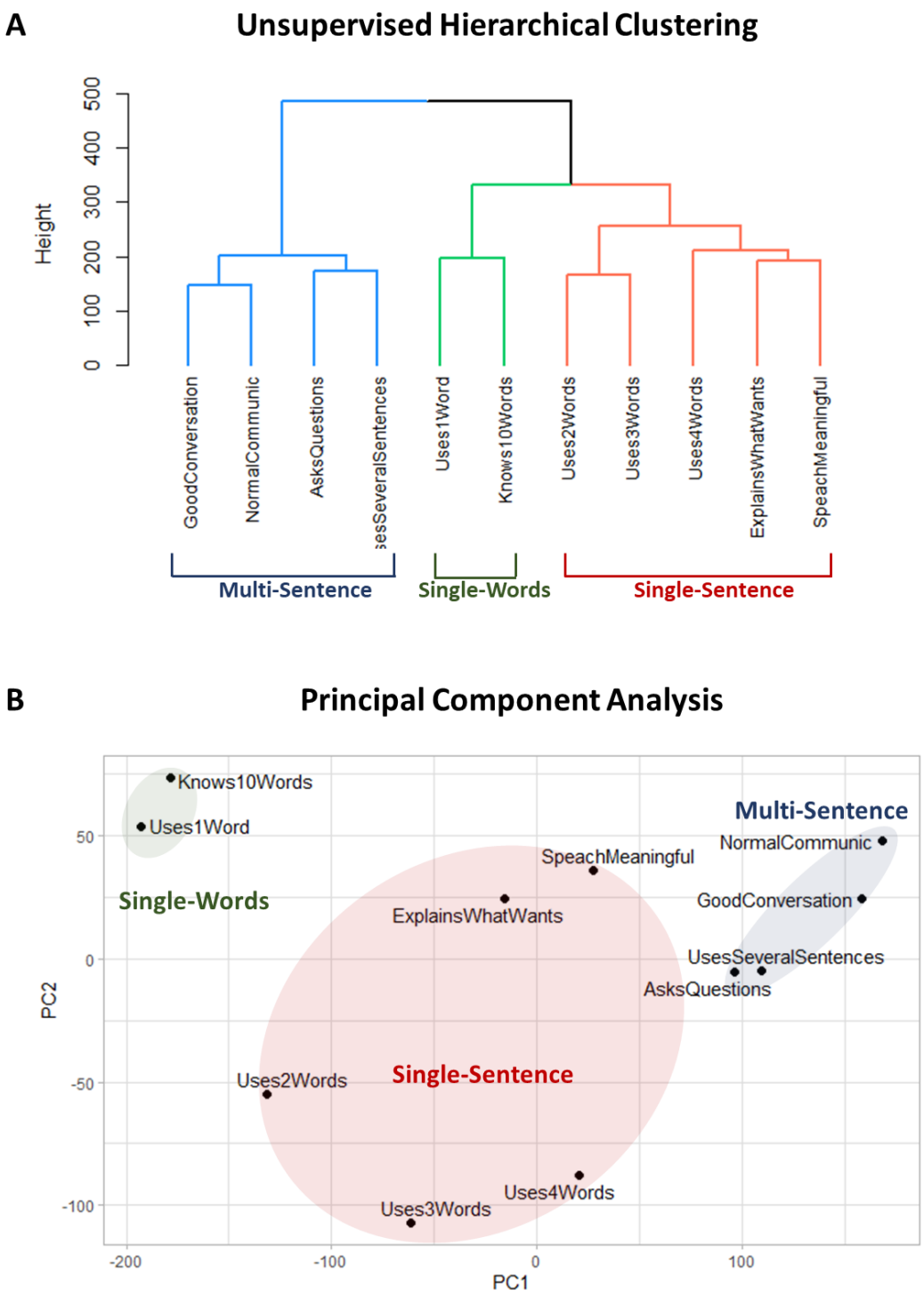

Figure 2. Clustering analysis of 11 expressive language items. (A) The dendrogram representing the unsupervised hierarchical clustering analysis of expressive language abilities. (B) Principal component analysis of expressive language abilities shows a clear separation between Single-Word, Single-Sentence, and Multi-Sentence items. Principal component 1 accounts for 51.2% of the variance in the data. Principal component 2 accounts for 10.9% of the variance in the data.

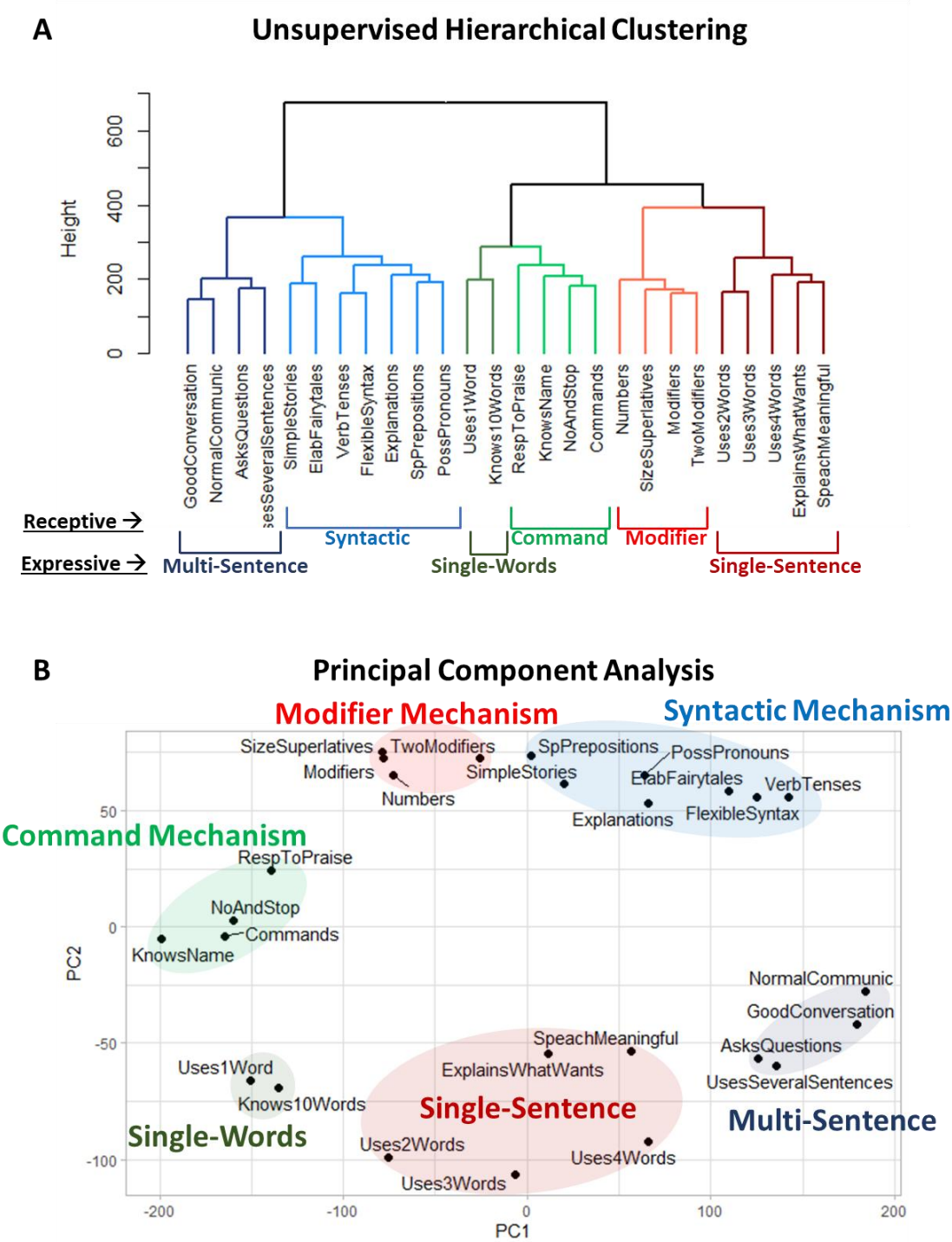

Figure 3. Clustering analysis of 15 receptive language and 11 expressive language items. (A) The dendrogram represents the UHCA of 26 items. (B) PCA of the 26 items shows a clear separation between the six clusters: three language comprehension clusters on the top (Command, Modifier, and Syntactic) and three expressive clusters on the bottom (Single-Word, Single-Sentence, and Multi-Sentence). Principal component 1 accounts for 37.3% of the variance in the data. Principal component 2 accounts for 11.0% of the variance in the data.

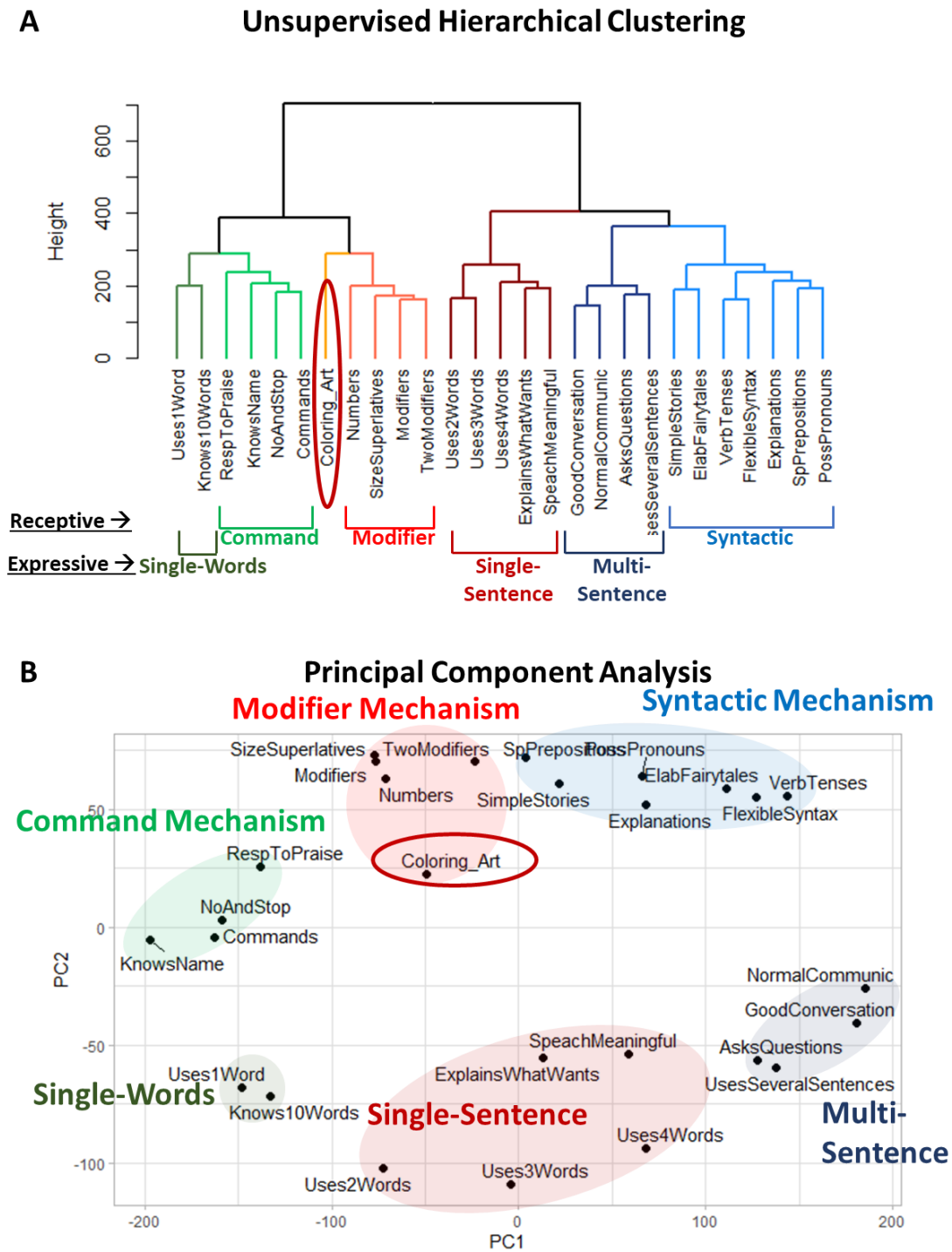

Figure 4. Clustering analysis of 15 receptive language and 11 expressive language items, along with the caregiver-reported item "[My child] does drawing, coloring, art" (labeled *Coloring_Art*). (A) Dendrogram generated using UHCA. (B) PCA, with Principal Component 1 accounting for 36.0% of the variance in the data and Principal Component 2 accounting for 10.6%. The *Coloring_Art* item clusters with the Modifier Mechanism.

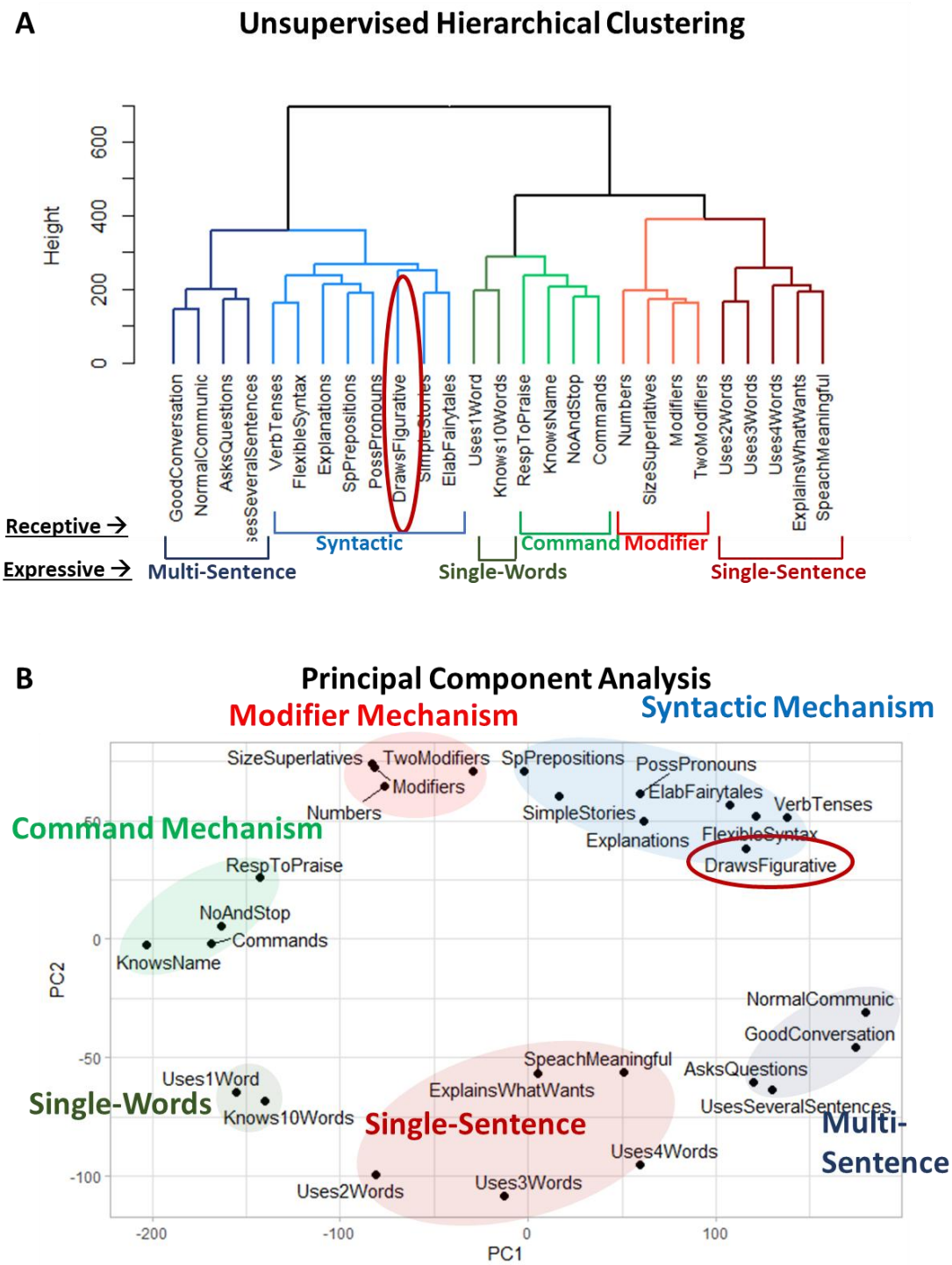

Figure 5. Clustering analysis of 15 receptive language and 11 expressive language items along with the caregiver-reported item "[My child] draws a VARIETY of RECOGNIZABLE images (objects, people, animals, etc.)" labeled *DrawsFigurative*. (A) Dendrogram generated using UHCA. (B) PCA, with Principal Component 1 accounting for 36.9% of the variance in the data and Principal Component 2 accounting for 10.6%. The *DrawsFigurative* item clusters with the Syntactic Mechanism.

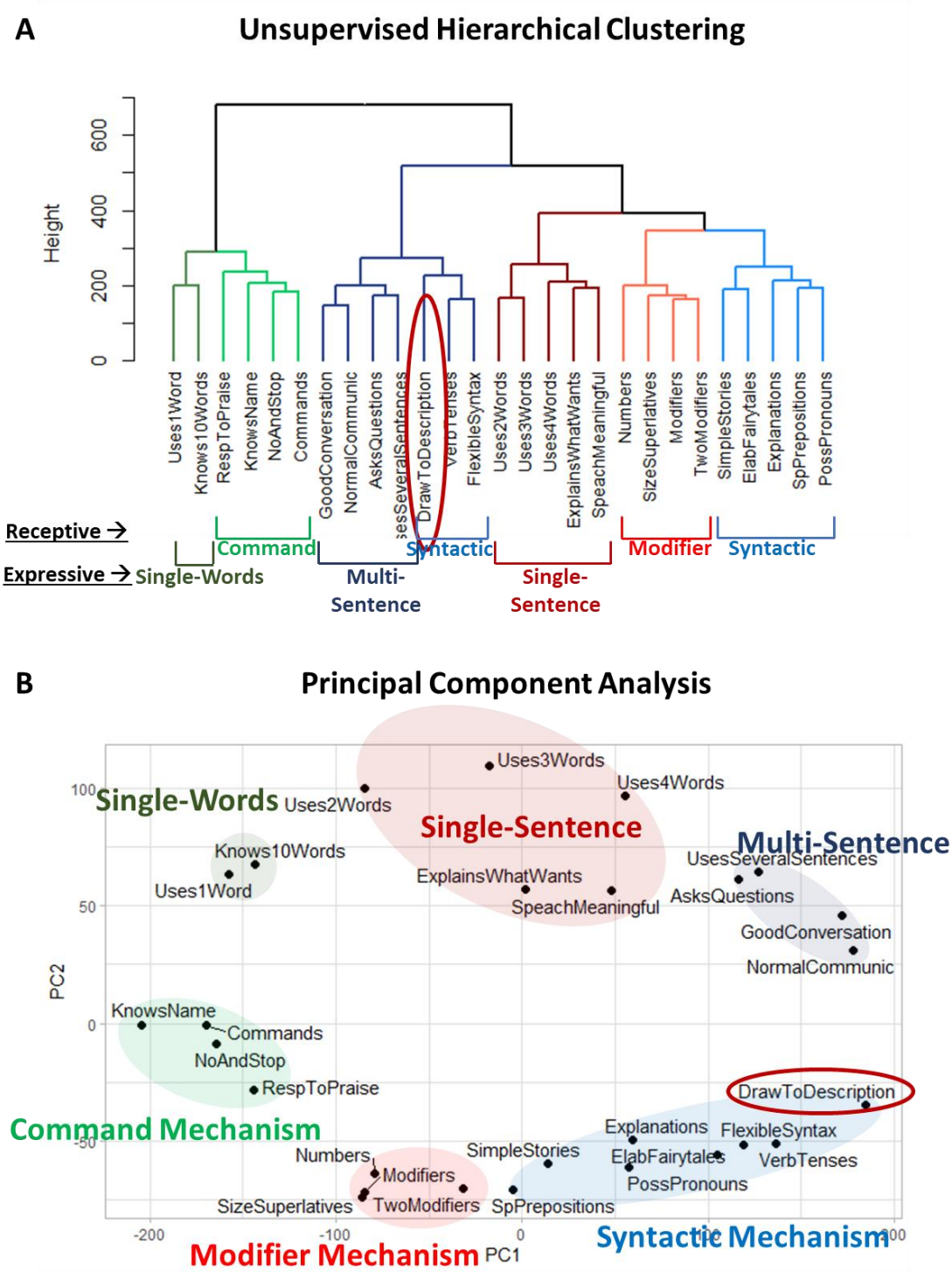

Figure 6. Clustering analysis of 15 receptive language and 11 expressive language items along with the caregiver-reported item "[My child] can draw a NOVEL image following YOUR description (e.g. a three-headed horse)" labeled *DrawToDescritption*. (A) Dendrogram generated using UHCA. (B) PCA, with Principal component 1 accounting for 38.5% of the variance in the data and Principal component 2 accounting for 10.5%. The *DrawToDescritption* item clusters with the Syntactic Mechanism.