International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Portability, Ecology and Adaptation: A Framework for Deploying Vocabulary Assessment Tools in Tribal Language Contexts in India

Dr. Stuti Bhagat

Jawaharlal Nehru University, New Delhi

Abstract

Standardised vocabulary assessment tools such as the Renfrew Word Finding Vocabulary Test (WFVT), the Peabody Picture Vocabulary Test (PPVT), and the MacArthur-Bates Communicative Development Inventories (CDI) are core instruments in developmental language research and clinical practice. The usage of wordlists as part of the methodology in the tribal language context in India raises fundamental questions of validity. The paper identifies three dimensions in which portability concerns arise: the item dimension, the norming dimension, and the construct dimension. A four-type taxonomy of item-level portability failure is proposed, distinguishing referent absence, referent mismatch, loanword substitution, and register mismatch. The paper proposes a three-stage adaptation framework as practical methodological guidance for researchers working in tribal language contexts comprising universality audit, ethnographic replacement, and community-referenced validation.

Keywords: Vocabulary Assessment, Portability, Tribal Languages, Kurukh, Linguistic Adaptation, Lexical Ecology, Wordlist Methodology

1. Introduction

Standardised vocabulary assessment tools occupy a central place in language research and clinical assessment. When researchers reach for the Renfrew WFVT, the PPVT, or the MacArthur-Bates CDI, they draw on decades of psychometric refinement and a substantial body of published work. This is, in itself, a sound and productive research practice.

What tends to go unexamined is the question of portability — whether an instrument designed for a particular population within a particular linguistic and cultural context can be carried into a new context and still be understood to measure the same construct. Portability is rarely treated as a methodological step in its own right. It is more often assumed: either assumed to hold, or assumed to be resolved by translation alone.

This paper argues that the assumption of portability deserves to be made explicit and tested. Wordlist adaptation — the process of examining and, where necessary, revising an instrument's items, norms, and procedural assumptions before deploying it in a new context — should be treated as a standard stage of research design. The argument presented in the current paper is that the existing tools are not universally inadequate, but that researchers should be explicit about the following points:

- where and why adaptation is needed,
- what form the adapted form takes, and
- how adapted instruments relate to their source versions.

To ground this argument, primary illustrative examples are drawn from Kurukh, a Dravidian language spoken by the Oraon tribe. Despite globalisation and various socio-economic developments, the tribal context retains its ecological, cultural, and linguistic uniqueness. The Kurukh reference illuminates the application of the vocabulary assessment tools in the case of tribal language contexts in India. Hence, the difference in the populations for whom standard assessment tools were designed and the referred communities of the current paper is substantial enough to render the relevant dimensions of adaptation highly visible. The paper also situates this argument within the history of wordlist methodology in Indian linguistic studies, from Grierson's Linguistic Survey of India through the Swadesh and Leipzig-Jakarta traditions, as these traditions have already developed principled tools for determining what travels across languages and what does not.

2. The Design Logic of Major Vocabulary Assessment Tools

2.1 The Renfrew WFVT

The Renfrew WFVT was developed in the late 1960s by Catherine Renfrew, a British speech therapist working with children in southern England (Renfrew). The design logic of the assessment tool is the picture-naming method in which children name the item presented to them, and their responses serve as a measure of expressive lexical retrieval. The item pool was constructed based on word frequency in British English child-directed speech of that period. The normative data were derived entirely from British English-speaking children.

When the Renfrew was administered to Indian and White English-speaking children in Durban, both groups scored significantly below the British norms, leading the authors to conclude the test was unsuitable for use with these populations in its present form (Pahl and Kara 69–73). Such studies foreground that the item selection is theoretically motivated and culturally specific — and these are not the same claim. An item appears frequently in British child-directed speech because it belongs to the domestic environment that British children inhabit. The frequency of a word is an intrinsic property of both the item and the context. Evidence from the existing cross-linguistic research supports that the environment (as the context) is a substantial factor in early vocabulary acquisition. For instance, Mandarin-speaking children acquire relatively more verbs and fewer nouns than English-speaking children at equivalent ages, a pattern traceable to differences in caregiver speech rather than any universal developmental sequence (Tardif et al. 620–635).

2.2 The Peabody Picture Vocabulary Test

The PPVT, developed by Lloyd Dunn in 1959, measures receptive vocabulary. It involved the child's ability to identify a pictured referent upon hearing a word aloud (Dunn). Within mid-century American psychology, the breadth of receptive vocabulary was widely understood as a proxy for verbal intelligence, and the test was shaped accordingly. Vocabulary breadth is a reflection of cognitive capacity and equally reflects which lexical domains a child needs to encounter. When a child with strong general cognitive ability has limited exposure to a specific item, it leads to poor performance. The poor performance reflects the differences in linguistic ecology, not a cognitive limitation.

Evidence of this problem in an Indian context comes from the Young Lives longitudinal study, which administered the PPVT to children across four countries, including India, where the test was rendered in

Telugu. Translation frequently distorted both the meaning and the relative difficulty of individual items. Where no single-word Telugu equivalent existed for an English term, translators were compelled to substitute the single words with descriptive phrases. For instance, 'hurdling' became 'the boy who jumps over the fence.' Items rendered in this way no longer carry the cognitive equivalence with their originals. Their retrieval difficulty had shifted in ways that led children to score artificially high and reach the ceiling of the test (Revollo and Scott). What makes this finding particularly instructive is the context in which it appeared. Young Lives was a well-resourced, carefully monitored study in which translation was approached systematically and in good faith. The difficulties that emerged were not the result of careless practice but an outcome of structural adaptation. Although translation was conducted with great care, it cannot manufacture equivalence where none exists. Thus, the limitation is embedded in the instrument's foundational assumptions, not in the choices made by those who attempted to adapt it.

2.3 The MacArthur-Bates CDI

The MacArthur-Bates CDI emerged from the longitudinal acquisition research of Bates, Bretherton, and Snyder (1988), developed as a parent-report instrument for tracking early lexical and grammatical development in infants and toddlers. The CDI's item pool was derived empirically from diary documentation of the toddler's communicative environment maintained by the parents. The parents played the most crucial role as the first and most consistent observers of their children's early vocabulary development. Therefore, the item pool, based on the observation of the vocabulary development of American English-speaking children, represents the speakers' linguistic environment growing up in the late twentieth century and simultaneously encodes a particular developmental ecology as the norm.

The CDI advisory board has long acknowledged that cross-linguistic deployment requires adaptation rather than translation. This principle is well articulated by Dale as one of the foundational guidelines for CDI developers (Adaptations). A CDI adaptation for Hindi-speaking children was attempted but ultimately abandoned as unsatisfactory (Dale and Penfold). Rasheed et al., adapting the CDI for rural Sindhi-speaking children in Pakistan, found that 158 of the original 258 words had to be dropped or replaced before the tool could function in the local context — and even then, the authors acknowledged that their findings were not generalisable "to urban or more advantaged settings" (431).

2.4 The Shared Hidden Assumption

The three assessment tools – the Renfrew WFVT, the PPVT, and the MacArthur-Bates CDI are comparatively distinct. However, they share one foundational assumption, which becomes visible in the tribal language contexts, i.e., the words a child is most likely to acquire early are drawn from an indoor, domesticated, consumer household world. The developmental sequence this environment produces is not biological or universally fixed. The early vocabulary of a child is overwhelmingly shaped by the interiors of their habitat, which is ecologically and culturally ordered. Thus, a different ecology will produce a different sequence. To use these tools unreflectively in tribal contexts is not simply to accept some tolerable degree of measurement error. It is to hold children to assumptions about what childhood looks like that are cultural rather than universal.

3. Three Dimensions of Portability Consideration

When a vocabulary assessment tool is deployed in a new language context without principled adaptation, it is imperative to identify what the instrument is actually measuring. It is a question that genuinely becomes difficult to answer. A significant ecological, cultural, or linguistic distance between the tool's developmental reference group and the community under scrutiny is inevitable. Thus, the three distinct

dimensions are proposed, which require careful examination, and each requires a qualitatively different kind of analytical attention. The discussion develops these through the lens of tribal language contexts in India, with particular reference to Kurukh.

3.1 The Item Dimension

Item-level portability refers to the specific elicitation stimuli that can reasonably be expected to function equivalently in a new context, wherein the item is emphatically not reducible to translation. Even when an item may carry a perfectly translatable label, the underlying referent plays little or no role in the target community's material world. Here, a functional diagnostic taxonomy is proposed based on the four recognisable patterns that item-level portability failures tend to cluster into:

Type 1 — Referent absence. The object depicted or named does not exist within the child's material environment. Items such as a pushchair may fall into this type. For many children in tribal communities, objects are absent from the child's ecological world altogether. To score such an item as a vocabulary failure is to commit a category error: the instrument is not measuring what it purports to measure.

Type 2 — Referent mismatch. The concept exists within the child's world, but the specific instantiation depicted does not align with their lived experience. Pahl and Kara identified the suitcase as an item on which South African Indian children performed poorly, suggesting the instrument's pictorial conventions encode objects from a specific British material world (71–72). Children in Kurukh-speaking communities shown a suitcase stimulus tended to identify the object as a box or bag rather than produce the target word, not because they lacked vocabulary, but because the depicted object did not correspond to the forms of luggage present in their environment.

Type 3 — Loanword substitution. The concept is familiar, and a word exists for it, but the word the child produces is a borrowing from a contact language. In Kurukh, for instance, there is no indigenous term for a bus. The English word is directly borrowed into everyday speech. A child who uses the word fluently may nevertheless be assessed as lacking the relevant vocabulary item if the scoring protocol does not account for loanwords as legitimate lexical knowledge.

Type 4 — Register mismatch. A native term exists for the referent, the child is familiar with the concept, and yet the word belongs to adult, specialist, or ritual vocabulary rather than to the everyday developmental lexicon of young children. Identifying specific instances in Kurukh would require systematic community piloting — a step identified here as necessary prior to any formal deployment.

3.2 The Norming Dimension

Even where individual items survive ecological scrutiny, a second and equally significant dimension of portability concerns the norms against which scores are interpreted. The normative data underpinning the Renfrew WFVT, the PPVT, and the MacArthur-Bates CDI were each derived from specific, historically and geographically bounded populations — British children in the case of the Renfrew, American children in the case of the PPVT and CDI. These norms are applied to interpret scores from a substantially different population, introducing a layer of interpretive uncertainty that should not be quietly absorbed into the analysis. It should be named, examined, and documented as part of the methodological record.

A Kurukh-speaking child who scores below the British norm for her age may be showing entirely typical development within her own community — or she may genuinely be showing early signs of atypical development. Without community-referenced norms, there is no principled basis for distinguishing between these two interpretations. Pahl and Kara found that when Renfrew norms were applied to both South African Indian and White English-speaking children in Durban, both groups scored significantly below the British age-group norms (69–73). The finding was not evidence of widespread language delay

among Durban's children. It was evidence that norms derived from one population do not automatically describe what typical development looks like in another.

3.3 The Construct Dimension

The third dimension of portability concerns the underlying construct that the instrument is designed to measure. The Renfrew WFVT purports to measure expressive lexical retrieval. The PPVT purports to measure receptive vocabulary breadth. The CDI purports to capture early lexical and grammatical development through the medium of parent report. These constructs are not in themselves culturally bounded. There is nothing culturally specific about the observation that children vary in their capacity to retrieve and produce words, or that some children command broader vocabularies than others. At this level of abstraction, the constructs travel reasonably well across contexts.

However, this dimension requires more careful examination of whether the assessment procedure actually presents the same cognitive task to children across different linguistic settings. A monolingual British English-speaking child completing the Renfrew WFVT is performing a lexical retrieval task in her only available language. A Kurukh-speaking child in Jharkhand who also commands Hindi and is assessed in either language is doing something meaningfully different. The child must not only retrieve a word but select and manage across the languages available to her. In both cases, the surface procedure looks identical, but the cognitive situation underlying shows disparity. Therefore, addressing the construct dimension properly requires explicit theoretical engagement, which is a willingness to articulate what the construct means in the specific research setting and what claims can and cannot legitimately be made on the basis of the scores that result.

4. Illustrative Evidence from Kurukh

4.1 Language, Community, and Ecological Context

Kurukh is a Dravidian language spoken by the Oraon tribe, principally across Jharkhand, Odisha, Chhattisgarh, and West Bengal, with smaller communities in Bangladesh and Nepal. The way of life of the Oraon community has historically been organised around settled agriculture, forest use, cattle rearing, and a richly elaborated cycle of communal ritual practice (Hahn; Roy).

Kurukh is spoken in a sociolinguistic context of active and daily multilingualism. Most Kurukh speakers command Hindi and at least one regional language — Nagpuri or Sadri, typically — with varying degrees of proficiency across domains and registers. The patterns of language socialisation within Kurukh-speaking communities are not uniform. While many children acquire Kurukh as their primary home and community language, others — particularly in more contact-intensive or migrant settings — grow up with Hindi or Sadri as their first language, with Kurukh acquired subsequently as part of a broader multilingual repertoire (Bhagat 26–36). Kurukh may therefore occupy the position of first, second, or co-acquired language depending on the specific household, locality, and social network in which a child develops. In India's tribal context, multilingualism is not a complicating variable — it is the norm. Any vocabulary assessment that fails to account for it has already misunderstood the context it is working in.

The lexical ecology of Kurukh reflects the physical and social world within which the language has developed. Kinship terminology is considerably more elaborated than in standard Indo-Aryan languages or English. Agricultural vocabulary encompasses terminology for soil types, cultivation stages, tools, grain varieties, and seasonal weather patterns. Forest ecology constitutes another semantically rich domain representing culturally and economically significant knowledge transmitted across generations. These are

the areas in which Kurukh-speaking children build their earliest, most secure vocabulary — and also the areas most thoroughly absent from existing assessment instruments.

4.2 A Lexical Audit of Renfrew Items Against Kurukh

Table 1 presents a preliminary illustrative audit of WFVT, PVT and CDI item types set against Kurukh lexical data, organised according to the four-type taxonomy proposed in Section 3.1. The items listed below are representative rather than exhaustive. The classifications are based on the author's observational knowledge as a trained linguist and as a member of the Kurukh-speaking community.

Table 1. Preliminary Lexical Audit of Renfrew WFVT Items Against Kurukh

Tool	Item	Kurukh Status	Portability Consideration
WFVT	Letterbox	No referent in community material world; no native term	Type 1 — Referent absence
WFVT	Stethoscope	Low ecological salience; term borrowed from Hindi/English if known	Type 1 — Referent absence
WFVT	Suitcase	Concept of carrying bag exists; Western hard-sided suitcase visually unfamiliar	Type 2 — Referent mismatch
WFVT	Pushchair	No cultural equivalent; babies carried rather than pushed in wheeled chairs	Type 2 — Referent mismatch
WFVT	Bus	Known object; term almost certainly a Hindi/English loan	Type 3 — Loanword substitution
WFVT	Plough	Native term exists; primarily adult/specialist agricultural vocabulary	Type 4 — Register mismatch
PPVT	Closet	No direct equivalent; storage organised differently	Type 1 — Referent absent
PPVT	Hurdling	No native term; concept culturally unfamiliar	Type 1 — Referent absent
CDI	Stove	Cooking done over open fire; known concept; borrow word from English/Hindi	Type 3 — Loanword substitution

Several observations follow. First, the portability considerations surfaced are not uniform in kind — they represent qualitatively distinct problems that call for qualitatively distinct responses. Type 1 items require outright replacement; Type 2 items may need only revision of the pictorial stimulus; Type 3 items require a prior decision about whether a loanword response constitutes valid Kurukh vocabulary knowledge; Type 4 items require community-referenced developmental data. Translation addresses none of these problems adequately and addresses some of them not at all.

Second, the areas in which Kurukh-speaking children would be likely to demonstrate the most confident expressive vocabulary are kinship terms, agricultural vocabulary, forest ecology terminology. They are almost absent not only from the Renfrew item pool but from the PPVT and CDI as well. The absence of the culturally primary vocabulary from an instrument is not visible unless someone looks for it — and it cannot be remedied unless that looking is done deliberately, with community knowledge and linguistic expertise as its guide.

Third, the portability classifications are not fixed properties of individual items. An item can present different portability problems depending on the child's ecological context. The pushchair illustrates this within the Renfrew: for a rural child, it represents a genuine referent absence; for an urban child familiar with the object, the same item becomes a loanword substitution problem. The CDI item stove makes the same point across the rural-urban divide: rural Kurukh children accustomed to cooking over an open fire encounter a referent mismatch, while urban children familiar with a stove borrow the term from Hindi or English, making it a loanword substitution problem. These observations are offered as hypotheses pointing toward a clear direction for future empirical work.

5. Wordlists in Indian Linguistic Studies: Methodological Precedents

There is a long and informative tradition of using standardised wordlists to survey Indian languages. Grierson's Linguistic Survey of India (1898–1928) applied wordlist-based methodology across 179 languages and 544 dialects (Grierson). There were limitations of Grierson's methodology for documenting languages. The data collection relied on institutionally affiliated intermediaries, who tended to record formal registers rather than everyday community speech. Transliterations were handled inconsistently across contributors, making systematic comparison difficult even within the survey itself (Majeed). An assumption is quietly embedded and less frequently discussed in the case of Grierson's design of language documentation is that every language will have a discrete native term for each listed item, and that the conceptual categories on the list would be intelligible and applicable across languages.

Morris Swadesh was directly concerned with addressing the problem of universal translatability. His contribution was to identify semantic categories that could plausibly be universal, occurring across cultures without distortion. These include body parts, basic kinship terms, and elementary natural phenomena. The items are based on the idea that these concepts are resistant to borrowing and culturally stable. If Swadesh's universality criterion were applied retrospectively to the Renfrew or PPVT item pools, a substantial proportion of their items would be flagged as culturally contingent rather than universal.

The Leipzig-Jakarta list extends the above concerns about documentation more rigorously. The list was developed based on cross-linguistic data. 100 items were derived from across 41 typologically diverse languages. They were selected specifically for their universal distribution and resistance to borrowing (Haspelmath and Tadmor). Notably, only 62 of its items overlap with Swadesh's 100-item list. It is a gap that reveals how many concepts Swadesh assumed to be universal were in fact more variable than he

recognised. In the framework proposed here, the Leipzig-Jakarta list serves as a first-pass filter: items that correspond to its concepts are provisionally retained or substituted.

6. A Framework for Principled Adaptation

The analysis presented above suggested that the vocabulary assessment tools adapted in the context of tribal language in India are not an optional methodological refinement but a substantive requirement. The three-stage framework proposed, which is a universality audit, ethnographic replacement, and community-referenced validation, is an attempt to provide a systematic structure in form without being rigid. The stages are laid out sequentially for clarity, though in practice the process is likely to be iterative.

Stage 1: Universality Audit

Before any data collection begins, every item in the assessment tool's existing pool needs to be examined individually. The purpose is not to condemn the above-mentioned instrument but to systematically establish that items can reasonably be expected to function in the new linguistic and cultural context, and which cannot. This examination involves two-level checks, which are at a theoretical level and at a community level. It is imperative to implement these checks before the researcher moves forward.

The theoretical check works by placing each item against the Leipzig-Jakarta list (Haspelmath and Tadmor). Items that correspond to Leipzig-Jakarta concepts, which are body parts, basic natural phenomena, and fundamental kinship terms, are provisionally retained. This will not imply that they are automatically usable. The pictorial stimulus may still need revision, and the scoring criteria may need adjustment. But they have at least cleared the first hurdle of cross-linguistic stability. Items that fall outside the Leipzig-Jakarta inventory are in a different position. Their presence on the original tool reflects the material world of the population for whom it was designed, not any claim to universality. This reflects the logic of standardisation rather than a limitation unique to any particular instrument. It does mean, however, that such items rest on culturally specific rather than universal foundations, and that an assumption requires examination before the items are carried into a new context.

The community level check is where the theoretical filter meets lived reality. A small group of adult community members, such as parents and primary caregivers of children in the relevant age range, is the most useful informant. They are asked to go through the flagged items one by one. The questions are straightforward:

- Is this object or concept part of their everyday life?
- Would a child of a particular age ordinarily come across it?
- Is there a word for a concept in the language that a young child would actually use?

These are some basic questions that require a cross-verification from people who know the community from the inside. A researcher working from outside, however carefully, cannot reliably answer them alone. Thus, local knowledge is irreplaceable. Items that fail either check are set aside for replacement, and the failure type drawn from the taxonomy in Section 3.1 is recorded against each one. This record is not just an administrative step. It is the documentation that allows the adapted instrument to be understood, evaluated, and, if necessary, challenged by other researchers working with the same communities later.

Stage 2: Ethnographic Replacement

Replacing a flagged item is not simply a matter of finding a local equivalent and adding it to the list. The substitute undergoes a verification through the criteria mentioned above for any assessment item to be considered. It needs to be something young speakers in this community encounter early, have a word for, and can recognise in a picture without ambiguity. This is the same rationale that any assessment item must

use. Extended fieldwork experience may attune a researcher to broad cultural patterns, but be deficient in reliably surfacing the fine-grained lexical distinctions that determine whether a child will recognise and name an item correctly. Instruments adapted primarily based on the researcher's judgement carry this risk without always acknowledging it.

Two practical constraints shape what counts as a viable replacement. The item needs to be pictorially representable without ambiguity, and it needs a single, stable native-language term that children would actually use. If the most common word for the concept is borrowed from Hindi or English, the replacement will simply reproduce the loanword problem it was meant to solve. Thus, finding items that meet both criteria requires going into the community and listening. Three sources are most useful:

- naturalistic observation of child-directed speech in everyday settings;
- caregiver interview, through which parents identify the words their children produced earliest and what vocabulary they consider age-appropriate; and
- consultation with community members who combine cultural knowledge with linguistic training.

Both cultural knowledge and linguistic expertise have to work together to make a meaningful contribution. For a Kurukh-speaking community, the domains that emerge from community elicitation will vary by context but are likely to include kinship terminology and local ecological and agricultural vocabulary — areas of cultural significance that existing tools do not account for.

Stage 3: Community-Referenced Validation

Following Stages 1 and 2, the adapted instrument requires empirical validation before it can generate interpretable scores. Community-referenced validation comprises three components. First, pilot testing with a representative sample of children from the target community was administered by examiners who are either native speakers or have received systematic training from native speakers. Pilot testing establishes whether adapted items consistently elicit clear and scorable responses, whether pictorial stimuli are reliably recognised, and whether the ordering of items reflects a meaningful gradient of increasing developmental difficulty.

Second, trained speech-language pathologists are virtually absent from tribal language communities. It means the assessment will frequently be conducted by community members who will be specifically trained for the task. The scoring criteria must be specified with enough precision that a non-specialist examiner can apply them consistently. Dialect variation, loanword responses, and partial responses must not be treated as complications to be penalised but as normal features of the linguistic landscape. They are ordinary features of how language works in multilingual, orally transmitted communities, and the scoring criteria need to reflect it.

Third, community-referenced norms need to be established. This means collecting scores from a sample of typically developing children within the community, grouped by age, to create local reference points. These norms are not comparable to British or American normative samples and should not be treated as such. Their purpose is to help distinguish typical from atypical development within this particular community, which is ultimately the question that matters most for clinical and educational practice.

7. Implications and Limitations

7.1 What the Framework Preserves

It is important to be clear from the outset about the scope of the proposed framework. The underlying constructs targeted by major vocabulary assessment tools, which are expressive lexical retrieval, receptive vocabulary breadth, and early lexical development, are valid and transferable across linguistic and cultural

contexts. These are genuine psycholinguistic phenomena that are present in Kurukh-speaking communities as much as in the populations for whom the existing instruments were originally designed. What the framework challenges is the assumption that the specific elicitation items, normative comparisons, and procedural conventions of tools designed for Western monolingual populations constitute valid measures of these constructs in tribal language contexts. The constructs themselves are not in dispute, but what requires critical examination is whether the instruments, as currently designed, can reliably capture them in a different ecological and cultural setting. Therefore, this is not an argument against standardised vocabulary assessment. It is an argument for assessment that is standardised in a way that is meaningful for the populations to whom it is applied.

7.2 Cross-Linguistic Comparability and Structural Barriers

A legitimate concern is that extensive localisation may undermine cross-linguistic comparability. The framework addresses two needs: the cross-linguistic comparability (by keeping items core) and local ecological validity (by culturally calibrated items). For instance, by anchoring the retained item core to the Leipzig-Jakarta list, the concepts are demonstrably stable across 41 typologically diverse languages. The adapted instrument preserves a theoretically grounded cross-linguistic reference point while the culturally calibrated replacement items provide the ecological validity necessary for locally meaningful clinical use.

A recent survey of practising Indian speech-language pathologists reported that assessment practices are substantially shaped by the unavailability of developed or adapted tools for Indian preschool-aged children (Rasheeka et al.). It suggests that clinicians are acutely aware of the problem while lacking the institutional support to address it systematically. Three directions for future work are particularly pressing: systematic universality audit of the Renfrew WFVT, the PPVT, and the CDI for major tribal language groups, including Santali, Mundari, Ho, Gondi, and Kurukh. A naturalistic vocabulary sampling in tribal language communities to provide an empirical foundation for ethnographic replacement. The development of training and capacity-building initiatives that enable community members to participate substantively in assessment administration and validation, since, given the near-total absence of speech-language pathologists with competence in tribal languages, community-based assessment models represent the only realistic pathway to achieving meaningful scale.

8. Conclusion

The argument advanced in this paper is straightforward. When a vocabulary assessment instrument is deployed in a linguistic context that differs substantially from the one for which it was designed, researchers bear a methodological responsibility to examine and document whether the instrument's items, normative assumptions, and construct definitions transfer meaningfully to the new setting. This does not imply that the existing tools are deficient; instead, it is a basic act of methodological due diligence.

The proposal here is not that these tools should be abandoned, but that the question of portability should be treated as an explicit, documented, and auditable stage of research design rather than an assumption left unexamined. The case of tribal languages in India gives this argument with particular clarity and force. The ecological, cultural, and linguistic distance between the populations for whom standard assessment tools were designed and the communities discussed here is large enough to make all three dimensions of portability, item validity, normative comparability, and construct transferability unmistakably visible. The evidence from the Young Lives study in India and the South African Renfrew application confirms that these are not theoretical concerns. They produce concrete, measurable distortions in data, causing ceiling

effects, norming anomalies, and items that behave in ways their design logic would not predict. The Swadesh and Leipzig-Jakarta traditions demonstrate further that principled tools for determining what transfers across languages already exist. The problem is not the absence of a method. It is the inconsistency with which it has been applied.

The framework proposed in the current paper cannot replace the prior decision to raise the portability question at all. That decision is a matter of methodological awareness of recognising that a vocabulary assessment tool carries a structured set of assumptions about which words exist in a language, which referents are culturally significant, and whose knowledge the assessment is designed to capture. The history of wordlist methodology in India, from Grierson's Linguistic Survey through the Swadesh tradition, shows that this recognition has been reached before, and then lost, at repeated points. The framework proposed here is an attempt to break that pattern by placing the portability question at the beginning of the research process, where it can be addressed deliberately, rather than at the end, where its consequences have already shaped the data.

References

1. Bates, Elizabeth, Inge Bretherton, and Lynn Snyder. *From First Words to Grammar: Individual Differences and Dissociable Mechanisms.* Cambridge University Press, 1988.
2. Bhagat, Stuti. "Crisis of Ethnic Identity with Language Loss of Oraon Tribe." *Towards Excellence*, vol. 10, no. 1, 2018, pp. 26–36.
3. Coffey, C., et al. "Cultural Variation in Early Vocabulary Composition: Italian and English Infants." *Journal of Child Language*, 2024. [Volume, issue, and page numbers pending verification.]
4. Dale, Philip S., and Michelle Penfold. *Adaptations of the MacArthur-Bates CDI into Non-U.S. English Languages.* CDI Advisory Board, 2011, mb-cdi.stanford.edu/documents/AdaptationsSurvey7-5-11Web.pdf.
5. Dunn, Lloyd M. *Peabody Picture Vocabulary Test.* American Guidance Service, 1959.
6. Fenson, Larry, et al. *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual.* Singular Publishing Group, 1993.
7. Grierson, George A. *Linguistic Survey of India.* Vols. 1–11, Office of the Superintendent of Government Printing, 1903–1928.
8. Hahn, Ferdinand. *Kurukh Grammar.* Bengal Secretariat Press, 1911.
9. Haspelmath, Martin, and Uri Tadmor, editors. *Loanwords in the World's Languages: A Comparative Handbook.* De Gruyter Mouton, 2009.
10. Krylova, Yulia. "Contact-Induced Lexical Replacement in South Odisha: Evidence from the Swadesh Wordlist." *Journal of South Asian Linguistics*, 2021. [Volume, issue, and page numbers pending verification.]
11. Majeed, Javed. *Colonialism and Knowledge in Grierson's Linguistic Survey of India.* Routledge, 2019.
12. Pahl, J., and M. B. Kara. "The Renfrew Word Finding Scale: Application to the South African Context." *South African Journal of Communication Disorders*, vol. 39, no. 1, 1992, pp. 69–73, doi.org/10.4102/sajcd.v39i1.279.
13. Rasheed, Muhammad A., et al. "Adaptation of MacArthur Communicative Development Inventory in Rural Pakistan." *Child: Care, Health and Development*, vol. 43, no. 3, 2017, pp. 427–434, doi.org/10.1111/cch.12450.

14. Rasheeka, J., et al. "Assessment Practices in Indian Speech-Language Pathology: A Survey of Clinicians." *Indian Journal of Speech-Language Pathology*, 2024. [Volume, issue, and page numbers pending verification.]
15. Renfrew, Catherine. *The Renfrew Word Finding Vocabulary Test.* Speechmark Publishing, 1968.
16. Revollo, P. E., and D. Scott. *Cognitive and Achievement Tests in the Young Lives Study.* Young Lives Technical Note 53, University of Oxford, 2022, www.younglives.org.uk/sites/default/files/2022-05/YL-TN53%20March%202022.pdf.
17. Roy, Sarat Chandra. *The Oraons of Chota Nagpur: Their History, Economic Life and Social Organisation.* Brahmo Mission Press, 1915.
18. Tardif, Twila, Susan A. Gelman, and Fen Xu. "Putting the 'Noun Bias' in Context: A Comparison of English and Mandarin." *Child Development*, vol. 70, no. 3, 1999, pp. 620–635.