

KESAHAN PENGUKURAN DAN KEBOLEHPERCAYAAN INSTRUMEN UJIAN STRUKTUR DALAMAN KATA IMBUHAN (USD*Ai*) BAGI PENGUASAAN MORFOLOGI ARAB

Hasnurol HASHIM*¹

Kaseh ABU BAKAR²

Maheram AHMAD³

¹Bahagian Biasiswa dan Pembiayaan, Kementerian Pendidikan Malaysia

^{2,3}Fakulti Pengajian Islam, Universiti Kebangsaan Malaysia

¹hasnurolhashim@gmail.com*

²kaseh@ukm.edu.my

³maheram@ukm.edu.my

Manuscript received 22 September 2021

Manuscript accepted 15 June 2022

**Corresponding author*

<https://doi.org/10.33736/ils.3963.2022>

ABSTRAK

Morfologi bahasa Arab bersifat kompleks berbanding morfologi bahasa lain. Hal ini kerana sistem morfologinya berasaskan kepada kata akar dan pelbagai lapisan proses *derivasi* (perubahan bentuk kata). Penentuan kata akar daripada perkataan berimbuhan Arab antara kesukaran pelajar bahasa Arab lebih-lebih lagi jika struktur dalaman sesuatu perkataan mengalami banyak *derivasi*. Justeru, kajian ini melaporkan kesahihan konstruksi dan kebolehpercayaan instrumen rintis bagi mengukur kesukaran struktur dalaman kata (SDK) imbuhan bahasa Arab. Kajian ini menggunakan pendekatan kuantitatif. Ujian Struktur Dalaman Kata Imbuhan (USD*Ai*) mengandungi 130 *item* aneka pilihan yang dibentuk berdasarkan teori SDK dan teori paradigmatis serta konsep derivasi bahasa Arab. Pemilihan sampel kajian menggunakan pensampelan bertujuan bersasar. Seramai 45 orang pelajar dari Kolej Universiti Islam Antarabangsa Selangor (KUIS) dipilih sebagai responden. Data dianalisis menggunakan Model Pengukuran Rasch. Nilai indeks kebolehpercayaan *person* ialah 0.95 ($\alpha = 0.97$) dan indeks pengasingan pula ialah 4.33. *Item* pula menunjukkan nilai kebolehpercayaan 0.87, indeks pengasingan 2.65 dan *mean infit* 0.97. Dapatan kajian menunjukkan 130 *item* USD*Ai* produktif bagi mengukur struktur dalaman kata (SDK) imbuhan bahasa Arab. Kajian ini turut memberi implikasi terhadap pembelajaran

morfologi Arab agar penguasaan struktur dalaman kata perlu dikuasai oleh para pelajar.

Kata kunci: Pengujian bahasa; morfologi; bahasa Arab; struktur dalaman kata; Model Pengukuran Rasch

MEASUREMENT VALIDITY AND RELIABILITY OF THE INTERNAL STRUCTURE OF THE DERIVATIONAL WORD INSTRUMENT FOR THE MASTERY OF ARABIC MORPHOLOGY

ABSTRACT

The Arabic morphology is more complex than that of other languages. This is attributed to the fact that Arabic morphology system is based on root and multi-layered derivational processes. The identification of root letters in derivational words constitutes one of the difficulties for Arabic learners. Therefore, this study reports the construct validity and reliability of a pilot instrument for measuring the difficulty of internal structure of Arabic derivational words. This study uses quantitative approaches. Internal Structure of Arabic Derivational Word Test (USDAl) consists of 130 *items* designed based on the Internal Structure of Word Theory, Paradigmatic Theory and the concept of Arabic derivation. Purposive sampling was employed to select the sample. 45 Arabic learners from Selangor International Islamic University College (KUIS) were selected as respondents. Data were analyzed using Rasch Model. Person reliability is 0.95 ($\alpha=0.97$) and person separation is 4.33. Item reliability is 0.87, item separation is 2.65 and mean infit is 0.97. All 130 items in USDAl are productive for measuring the internal structure of Arabic derivational words. This study also has implications for the overall learning of Arabic morphology so that the mastery of the internal structure of the word needs to be mastered by students.

Keywords: Language testing; morphology; Arabic; internal structure of word; Rasch Model

Pengenalan

Pengujian bahasa telah mengalami perubahan evolusi dan revolusi dalam beberapa dekad yang lalu (Allan 1992; Alderson 1998; Alderson & Banerjee 2001; Cumming 1996; Hassan Basri, 2002; Maeda & Naoumi, 2004; Abu Bakar, 2008; Song, 2002). Perkara tersebut diperakui oleh Gregory et al. (2013), Perneger et al. (2014) dan Rahman (2016) bahawa pengujian bahasa pada masa kini memainkan peranan penting dalam penyelidikan bagi memperoleh sesuatu maklumat yang diperlukan. Dalam konteks pengujian standard untuk bahasa Arab sebagai bahasa kedua atau bahasa asing pula telah pun berkembang sejak tahun 2000 (Al-Najem, 2007). Hal ini berikutan

pengujian bahasa Arab semakin banyak digunakan di beberapa keadaan untuk membuat pelbagai keputusan, kajian pengesahan dan apa-apa jua pemeriksaan. Steven (2008) menyatakan pengujian berkaitan bahasa sama ada bahasa ibunda mahupun bahasa kedua, kini menjadi asas untuk menentukan prestasi atau pencapaian akademik sesuatu komuniti.

Seterusnya menurut Kusaeri (2012), penilaian mahupun pengukuran tidak boleh dipisahkan daripada proses pengajaran dan pembelajaran (PdP) pelajar. Perkara tersebut dinyatakan kerana menjadi pendorong kepada para guru untuk menggunakan strategi atau pendekatan pengajaran yang lebih baik dan bermotivasi. Al-Najem (2007) turut menegaskan orang ramai kini mempunyai akses ke pelbagai sumber maklumat mengenai cara pengukuran atau pengujian bahasa Arab disahkan dan tujuan ia harus digunakan serta prosedur yang dipandu untuk memastikan ketepatan skor ujian. Menurutnya lagi, tindakan sedemikian akan membantu mencapai profesionalisme bidang penilaian bahasa Arab.

Dalam konteks pengujian dan pengukuran bahasa Arab di Malaysia, terdapat kajian-kajian mengenainya tertumpu kepada penyelidikan morfologi Arab oleh sarjana lepas seperti Abu Bakar (2008), Jalil (2012), Rasad dan Mahamod (2013), Norazwa (2014), Baharudin et al. (2017), Asbulah et al. (2018), Maskor et al. (2018) dan Tuan Omar et al. (2018). Kebanyakan dapatan kajian melaporkan kerumitan pelajar terhadap penguasaan dan permasalahan morfologi terutamanya aspek kata mahupun kosa kata Arab. Hakikatnya, kepentingan aspek kata dan kosa kata ini memberi impak besar kepada penguasaan kosa kata dan kemahiran bahasa kedua. Diperakui oleh Al-Hafiz Abd Rahim (2006) dan Rosni (2009) bahawa pelajar bukan penutur jati Arab, seharusnya bermula dengan pembelajaran aspek kata dan strukturnya menerusi morfologi Arab. Kajian-kajian ini juga mengetengahkan peri penting mengukur penguasaan morfologi Arab dan mengkaji kesilapan yang dilakukan pelajar dalam bahasa. Namun kajian-kajian lepas yang dilakukan, kebanyakannya tidak bersandarkan teori yang kukuh dalam pembinaan sesuatu instrumen bagi mengukur dan menguji permasalahan aspek kata dalam bidang morfologi Arab.

Justeru menerusi perbincangan ini, penulis membina sebuah instrumen berdasarkan kerangka teoritikal dan konseptual kajian bagi mengenal pasti permasalahan berkaitan aspek kata dan kosa katanya kemudian mengesahkan faktor kesukaran struktur dalaman kata (SDK) imbuhan bahasa Arab. Instrumen yang dibina ini dikenali sebagai Ujian Struktur Dalaman Kata imbuhan (USDAl) bagi pelajar bahasa Arab. Pengaplikasian teori struktur dalaman kata (SDK) oleh Tucker (2011) dan Booij (2015), teori *paradigmatis* Bloomfield (1970) dengan konsep derivasi (*al-Isytiqaaq*) bahasa Arab kata imbuhan oleh Muhammad Qadur (1999) dan Maher Sya^cban (2011) dijadikan sandaran teori pembinaan instrumen USDAl. Di samping itu, pendekatan model pembinaan instrumen oleh McIntire dan Miller (2007) turut diadaptasikan sebagai panduan dalam proses pembinaan instrumen.

Perkara tersebut kerana bahasa Arab mempunyai sistem morfologi yang tinggi terhadap pembinaan kata dan strukturnya yang sentiasa mencirikan proses *derivasi* (perubahan bentuk kata). Hal demikian menghalang kepada penguasaan kosa katanya

serta perkembangan kemahiran bahasa. Sehubungan itu, kajian empirikal berteori ini mampu mengesahkan kesukaran kosa kata imbuhan mengikut hierarki kerumitan berdasarkan proses perubahan kata menerusi proses derivasi sama ada melalui proses pengimbuhan (*mazid*), pengguguran (*hazaf*), pengalihan (*ʿlāl*), penggantian (*ibdāl*) mahupun penyerapan (*idghām*). Selain itu, kajian ini juga dapat mengetahui lebih tepat serta pembuktian mengenai penguasaan pelajar terhadap struktur dalaman kata (SDK) imbuhan bahasa Arab.

Kesimpulannya, hasil perbincangan menunjukkan struktur dalaman kata (SDK) dalam bidang morfologi, seharusnya dikuasai dan dipelajari terlebih dahulu oleh kalangan pelajar bahasa Arab berbanding bidang-bidang linguistik bahasa yang lain. Hal ini berdasarkan implikasi kegagalan terhadap struktur dalaman kata (SDK) yang ditemui, memberi impak kepada penguasaan kosa kata dan kemahiran bahasa terutamanya bahasa Arab sebagai bahasa kedua di Malaysia.

Sorotan Kajian

Menurut Gregory et al. (2013) dan Rahman (2016), pengukuran dan pengujian dalam penyelidikan digunakan sebagai skala pemboleh ubah seperti lulus, memuaskan, baik atau gagal bagi mengukur ciri pelajar dari aspek pengetahuan, kemahiran, kebolehan, keupayaan, kecekapan dan minat mereka. Pengukuran dan pengujian juga berfungsi sebagai prasyarat yang diperlukan oleh penyelidik untuk penambahbaikan sesebuah instrumen atau protokol bagi memperoleh maklumat (Gregory et al., 2013; Perneger et al., 2014). Ia turut dipersetujui oleh Clapham (2000) yang menyatakan majoriti ahli bahasa menggunakan pendekatan pengukuran dan pengujian sebagai teknik menarik minat dan membantu mengumpulkan data penyelidikan mereka. Pendekatan tersebut digunakan untuk menganalisis data berkaitan pengetahuan mahupun penggunaan bahasa terutamanya dalam linguistik terapan.

Beberapa kajian terkini berkenaan pengujian kosa kata bahasa kedua atau asing telah dilakukan dan menghasilkan beberapa hasil dapatan yang unik. Kajian tersebut bukanlah perkara baru dilakukan pada peringkat global mahupun pada—peringkat Malaysia bahkan beberapa kajian mengenainya telah pun dilakukan. Justeru, tinjauan terhadap kajian-kajian lepas dilakukan bagi melihat kaedah dan metodologi sesuatu pengukuran atau pengujian dilakukan menepati prosedur yang dipersetujui oleh para sarjana terdahulu berkaitan pengukuran dan pengujian bahasa Arab sebagai bahasa kedua. Aspek kesahihan konstruksi dan kandungan serta kebolehpercayaan instrumen dititikberatkan bagi mencapai dapatan yang bersifat empirikal.

Menerusi kajian Al-Najem (2007) berkaitan pengukuran dan pengujian bahasa Arab sebagai bahasa kedua atau *Assessment of Arabic as a Foreign language (AFL)* dalam konteks pelajar di Amerika. Skala *Interagency Language Roundtable (ILR)* bagi program *Arabic as a Foreign language (AFL)* intensif digunakan untuk membuat pelbagai keputusan termasuklah yang berkaitan penempatan, kemasukan dan penyingkiran. Ujian yang digunakan mengukur tiga bahagian, iaitu pemahaman mendengar, pemahaman membaca dan kecekapan berbahasa. Dua jenis pengujian dilakukan, iaitu

Ujian Kemahiran Bahasa Arab (*Arabic Productive Test*) dan Ujian Bahasa Arab (*Arabic Skills Test*). Ujian-ujian tersebut digunakan bagi pencapaian yang digunakan sebagai kuiz pertengahan semester dan akhir semester termasuk peringkat permulaan, pertengahan dan lanjutan. Keberkesanan hasil kajian dapat memberikan banyak maklumat mengenai kebolehpercayaan dan kesahan skor ujian pelajar di samping mengembangkan bahasa Arab sebagai bahasa kedua atau bahasa asing di Amerika.

Kajian pengukuran atau pengujian bahasa Arab sebagai bahasa kedua turut ditemui seperti kajian yang dijalankan oleh Abu Bakar (2008) dengan mengemukakan kajian kesahan terhadap 134 *item* untuk mengenal pasti bacaan bahasa Arab dan mengkalibrasi *item* berdasarkan satu pemboleh ubah bagi kefahaman bacaan bahasa Arab. Hasil dicapai menerusi aplikasi Model Pengukuran Rasch pada reka bentuk dan kesahan *item* ujian. Pada akhir kajian, beliau menghasilkan sebuah *item* bank permulaan yang mengandungi *item* yang sedia dikalibrasi bagi kegunaan ujian diagnostik kemahiran bacaan bahasa Arab adaptasi berasaskan komputer. Kajian menerusi aplikasi Model Pengukuran Rasch ini turut dilakukan oleh Asbullah et al. (2018) bagi mengukur tahap pengetahuan kolokasi bahasa Arab dalam kalangan graduan bahasa Arab di universiti awam, Seramai 61 orang pelajar awam telah dijadikan responden dan dapatan kebolehpercayaan *item* menunjukkan tahap kebolehpercayaan yang tinggi, iaitu 0.91 manakala indeks kebolehpercayaan responden ialah 0.81. Beliau menyatakan Model Pengukuran Rasch telah menunjukkan instrumen Pengetahuan Kolokasi bahasa Arab (i-KAC) yang dibangunkan mempunyai tahap kesahan dan kebolehpercayaan yang tinggi.

Metodologi pengukuran dan pengujian menggunakan aplikasi Model Pengukuran Rasch ini turut dilakukan oleh Maskor et al. (2018). Hasil penganalisan menggunakan Model Pengukuran Rasch dapat mengukur nilai rata-rata infiniti dan pakaian persegi (MNSQ) serta memberi impak kebolehpercayaan pengujian yang memuaskan. Bagi nilai kebolehpercayaan berdasarkan *Cronbach Alpha* dengan julat nilai adalah sesuai. Kesahan dan kebolehpercayaan instrumen turut diperiksa dan diterima. Menurutnya, Model Pengukuran Rasch adalah sesuai bagi prosedur pengesahan instrumen dan kandungannya. Hasil kajian telah mengemukakan *Arabic Productive Vocabulary Knowledge Test* (APVKT) sebanyak 23-*item* yang boleh digunakan dalam penyelidikan dan pendidikan. Ketepatan hasil kesahan kandungan pengukuran dan pengujian diperteguhkan oleh Baharudin et al. (2017) bahawa hasil kajian terhadap pembinaan ujian dan penentuan tahap saiz kosa kata Arab menunjukkan nilai korelasi adalah tinggi, iaitu 0.84 manakala pekali *Cronbach Alpha* turut tinggi, iaitu 0.89 dan pekali *cohen kappa* juga dilihat tinggi, iaitu 0.95. Sehubungan itu, kesahan isi kandungan kajian mereka turut diteliti daripada semakan panel pakar.

Selain bahasa Arab, bahasa Inggeris juga sebagai bahasa kedua turut dilakukan kajian pengukuran dan pengujian seperti kajian yang dilakukan oleh Song (2002) yang mengemukakan keputusan dalam pengujian *Productive Level Test* (PLT) ke atas pelajar universiti di Jepun. Beliau melaporkan bahawa pengetahuan kosa kata yang produktif pelajar Jepun dan Cina pada tahap 2,000 perkataan (2,000 *word-level*) adalah tidak mencukupi untuk menghadapi tugas-tugas yang dikehendaki oleh kursus bahasa Inggeris sebagai bahasa kedua bagi pelajar Universiti. Hal ini diperkukuh menerusi kajian

yang dilakukan oleh Maeda dan Naoumi (2004) yang membangunkan soal selidik 35 perkara mengenai kosa kata Inggeris sebagai bahasa kedua sebagai strategi pembelajaran dan mengujinya di universiti Jepun. Hasil pengujian menunjukkan pelajar sering menunjukkan minat dalam membangunkan kosa kata, namun pada masa yang sama mereka tidak pasti tentang keperluan penguasaan kosa kata tersebut dilakukan.

Berdasarkan tinjauan kajian berkaitan pengukuran & pengujian bahasa kedua atau bahasa asing (*L2-second language testing*) di atas, dapat disimpulkan bahawa setiap pengukuran dan pengujian seharusnya menepati prosedur pengukuran yang standard terutamanya bagi kegunaan bahasa kedua seperti bahasa Arab. Prosedur yang dimaksudkan bermula dengan pembinaan instrumen yang berasaskan model pembinaan dan pengukuran yang menepati prosedur pengesahan instrumen menerusi teori tindak balas *item* (*item response theory*). Selain itu, kandungan instrumen termasuk *item* yang dibentuk melalui proses kesahan isi kandungan dan *item* ujian agar pengesahan dan tindak balas *item* yang dilakukan memberi kebolehpercayaan dan kesahan skor ujian pelajar. Sehubungan itu, tahap frekuensi penggunaan perkataan turut diteliti bagi menguji tahap penguasaan kosa kata dalam sesebuah bahasa mengikut peringkat pengajian. Hasil instrumen yang dibina pula dianalisis secara sistematik agar dapatan kajian dilaporkan mencapai kesahihan konstruksi dan kebolehpercayaan instrumen yang menepati tahap penguasaan atau pengetahuan para pelajar bahasa Arab.

Pembinaan Instrumen USDAi Permodelan Rasch

Kajian ini dilakukan menggunakan ujian bertulis aneka pilihan yang mengandungi 130 *item*. Pada peringkat awal, konstruksi ujian dibahagikan kepada lima dimensi bentuk kata atau jenis perkataan Arab, iaitu:

- (1) Kata akar *Sahih*,
- (2) Kata akar *Mu^ctal*,
- (3) Imbuhan *Sahih*,
- (4) Imbuhan *Mu^ctal*,
- (5) Imbangan pola atau *wazan*.

Dimensi tersebut ditakrifkan sebagai keupayaan mengenali atau menghasilkan pemahaman struktur dalam kata (SDK) termasuk pembentukan perkataan berpandukan penguasaan *wazan* atau imbangan pola. Pendekatan pensampelan perkataan (*sampling word*) yang dipilih pula adalah berdasarkan senarai kekerapan daripada *A Frequency Dictionary of Arabic* oleh Buckwalter dan Parkinson (2011) pada tahap penggunaan 3000 frekuensi perkataan. Perkataan yang dipilih kemudian diimbangi dan disesuaikan dengan perkataan yang terdapat dalam *al-Tasrifyyah* oleh *al-Syeikh Muhammad Ma'sum Ali* (1965) dan kamus *Tasrif al-Af'al* oleh Ghizawiy (2008). Sekiranya perkataan yang dipilih daripada Buckwalter dan Parkinson (2011) tidak terkandung dalam kamus *tasrif* tersebut, perkataan akan ditinggalkan dan digantikan dengan kata-kata lain yang terdapat dalam kedua-dua sumber tersebut. Pensampelan perkataan turut dilakukan secara sistematik yang merujuk nisbah 1:3 ke atas 180 perkataan yang terdiri daripada

kata kerja dan kata nama. Justeru, terdapat 60 *item* perkataan berdasarkan tahap 3000 frekuensi perkataan yang melibatkan kata akar dan imbuhan dalam sesuatu perkataan.

Setelah *item* disahkan oleh pakar, hanya 130 *item* diterima daripada 138 dalam instrumen yang dibangunkan. Instrumen ini menggunakan pendekatan pelbagai pilihan jawapan yang digunakan secara meluas untuk menilai hasil pembelajaran pelajar, iaitu mencakup penguasaan struktur dalaman kata (SDK) imbuhan bahasa Arab. Walaupun, kandungan instrumen termasuk *item* yang dibentuk melalui proses kesahan isi kandungan dan *item* ujian, namun pengesahan dan tindak balas *item* yang dilakukan memberi kebolehpercayaan dan kesahan skor ujian pelajar. Hal ini kerana instrumen yang *reliable* dan *valid* akan memberikan maklumat yang dipercayai dan diyakini. Sebaliknya, instrumen yang gagal memenuhi kedua syarat tersebut akan memberikan hasil yang *bias* sehingga meragui kualiti sesuatu kajian.

Hasil pembinaan instrumen pula memastikan dapat memberi banyak maklumat mengenai pencapaian, kemampuan atau pengetahuan responden yang menjalani proses kajian, seterusnya melaporkannya sebagai maklumat baharu dalam bidang penyelidikan. Justeru bagi mencapai tujuan tersebut, Model Pengukuran Rasch digunakan kerana ia diiktiraf secara meluas sebagai model respon *item* yang paling mudah (Aziz et al., 2017; Khoiry et al., 2019) dan mampu mentaksirkan keupayaan seseorang dalam bentuk statistik yang berguna (Azami et al., 2019) di samping menawarkan peluang yang sangat besar untuk siasatan kesahan. Oleh itu, kajian ini bertujuan:

- (1) Mengenal pasti indeks kebolehpercayaan instrumen menggunakan Model Pengukuran Rasch,
- (2) Mengenal pasti indeks pengasingan *item* ujian dan kecukupan *person* (responden),
- (3) Mengenal pasti *item* instrumen bergerak dalam satu arah dan menepati tujuan kajian.
- (4) Mengenal pasti *item* yang sesuai dalam pembinaan ujian dalam dimensi yang boleh diterima,
- (5) Mengenal pasti ciri unidimensionaliti atau analisis piawaian varians komponen utama (PCA).

Kaedah Kajian

Kajian ini menggunakan kaedah kuantitatif bersifat pensampelan bertujuan bagi melaporkan secara statistik mengenai tahap penguasaan dan kefahaman pelajar terhadap struktur dalaman kata (SDK), iaitu kata akar, huruf tambahan (imbuhan) dan *wazan* kata kerja imbuhan bahasa Arab dalam kalangan pelajar. Tatacara kutipan data dilakukan dengan pengedaran instrumen secara dalam talian dan anggaran masa menjawab ialah selama satu jam. Berikut elemen asas bagi tujuan pengumpulan data kajian:

(1) Instrumen

Ujian struktur dalaman kata 130-*item* Arab aneka pilihan jawapan yang dikenali sebagai Ujian Struktur Dalaman Kata Imbuan (USDAi) ditadbir kepada responden secara dalam talian. Mereka dikehendaki memilih jawapan yang terbaik untuk setiap *item* dengan pilihan A, B C atau D. Masa yang diperuntukkan untuk menjawab semua *item* ialah 60 minit atau satu jam.

(2) Peserta

Seramai 45 orang pelajar dari Kolej Universiti Islam Antarabangsa Selangor (KUIS) yang sedang mengikuti pengajian bahasa Arab dipilih bagi mewakili populasi berlatar belakang pengajian bahasa Arab sama ada pada peringkat tamhidi (kelas persediaan), ijazah sarjana muda dan pasca siswazah. Pengkaji menggunakan pensampelan bertujuan (*purposive sampling*) bagi menepati kriteria dan ciri yang dikehendaki pengkaji. Kriteria dan ciri pensampelan bertujuan tersebut merujuk kepada latar belakang pendidikan sama ada di sekolah agama atau mengikuti kurikulum pendidikan bahasa Arab Kementerian Pendidikan Malaysia (KPM). Selain itu, kriteria yang dirujuk juga berdasarkan penawaran mata pelajaran bahasa Arab sama ada major dan minor atau mengambil mata pelajaran bahasa Arab secara pilihan di universiti awam (UA). Seterusnya, tenaga pengajar yang terlibat dengan kelas pengajaran di universiti tersebut mencadangkan para pelajar yang memenuhi kriteria pengkaji dijadikan sebagai responden kajian.

Semua pelajar berumur 18 tahun hingga 27 tahun yang mempelajari bahasa Arab sejak berumur tujuh tahun atau daripada peringkat sekolah rendah kurikulum KPM dan mempunyai kepelbagaian kebolehan kognitif berdasarkan tahap pengajian semasa. Hal ini kerana menurut Rabiah Tul Adawiyah et al. (2020), pemerolehan morfologi bahasa kedua seharusnya dilihat daripada peringkat umur tujuh tahun yang telah mengikuti kurikulum sejak sekolah rendah lagi. Maka rasional kajian terhadap perkara tersebut adalah ingin mengenal pasti penguasaan struktur dalaman kata (SDK) imbuan yang merupakan perkara asas dalam pembelajaran bahasa Arab oleh kalangan pelajar di setiap peringkat pengajian. Selain itu, kajian ini bertujuan menguji *item* yang dibina dapat mengimbangi aras kognitif yang pelbagai. Berikut latar belakang sampel yang menjadi responden dalam kajian ini:

Jadual 1*Profil Demografi Kajian*

Jantina		Tahap Pengajian			
Lelaki	Perempuan	Peringkat Sijil	Peringkat Diploma	Peringkat Sarjana Muda	Peringkat Sarjana/PhD
30	15	6	5	32	2

Jadual 1 menunjukkan profil demografi responden berlatar belakang pengajian bahasa Arab yang terlibat dalam kajian ini. Dari segi komposisi jantina, data menunjukkan 30 responden ialah lelaki dan 15 orang ialah perempuan. Berdasarkan tahap pengajian menunjukkan bahawa responden yang berada di tahap sijil ialah seramai 6 orang, peringkat diploma ialah 5 orang, peringkat sarjana muda ialah 32 orang dan peringkat lanjutan sama ada peringkat sarjana atau Ph.D seramai 2 orang. Kepelembagaan peringkat tahap pengajian oleh responden di atas adalah bertujuan memastikan instrumen yang diedarkan menepati pengukuran tahap penguasaan sama ada pada kategori rendah, kategori sederhana mahupun kategori tinggi.

(3) Analisis Data

Analisa data dilakukan menggunakan Model Pengukuran Rasch untuk data dikotomi. Perisian yang digunakan ialah *Winsteps* versi 5.0.2.0 yang dibangunkan oleh Linacre (2021). Tujuan menggunakan Model Pengukuran Rasch ini bagi mengukur dan memastikan *item-item* yang dibina mempunyai kejituan dan bersifat tekal (Abu Bakar & Hassan, 2011). Model Pengukuran Rasch merupakan model respon *item* yang paling mudah (Aziz et al. 2017; Khoiry et al. 2019) dan mampu mentaksirkan keupayaan individu dan mengklaborasi *item* pada satu skala pengukuran (Azami et al., 2019) dan menawarkan peluang yang sangat besar untuk siasatan kesahan (Bond & Fox, 2015). Sehubungan itu, menurut Aziz et al. (2017), pengaplikasian Model Pengukuran Rasch dalam sesuatu kajian dapat membina skala berdasarkan set *item*. Oleh itu, Model Pengukuran Rasch sangat bersesuaian dengan kajian ini kerana dapat memberikan properti psikometrik bagi tujuan pengesanan USDAi sebagai skala pengukur dan paling penting ia dapat menunjukkan hierarki kesukaran *item* secara objektif dan statistik. Hal ini kerana kegunaan psikometrik tersebut disokong oleh hasil skala yang digunakan (Ullah et al., 2021).

Dapatan Kajian

Berikut Jadual 2 mengenai rumusan analisis indeks kebolehpercayaan dan pengasingan *person* dan *item* bagi kajian ini:

Jadual 2

Rumusan Analisis Kebolehpercayaan Person dan Item

<i>Person</i> (r = 45)		
Indeks pengasingan <i>person</i>	4.33	
Indeks kebolehpercayaan <i>person</i>	0.95	
Min (Purata)	1.51	
Maks	5.7	
Min	-1.58	<i>Cronbach alpha</i>
<i>Item</i> (i = 130)		= 0.97
Indeks pengasingan <i>item</i>	2.65	
Indeks kebolehpercayaan <i>item</i>	0.87	
Min (Purata)	0.00	
Maks	3.25	
Min	-3.07	

(1) Indeks Kebolehpercayaan

Jadual 2 di atas memaparkan ringkasan statistik untuk *person* (penguji ujian/responden) dan *item* (soalan USDAi). Secara spesifik Model Pengukuran Rasch ini mengukur dua indeks kebolehpercayaan, iaitu indeks kebolehpercayaan *person* dan indeks kebolehpercayaan *item*. Indeks kebolehpercayaan *person* merujuk andaian kebolehan individu dalam sampel adalah tekal walaupun diberikan set yang lain tetapi mengukur konstruksi yang sama. Manakala indeks kebolehpercayaan *item* merujuk kepada kesukaran *item* adalah sama berbanding sampel lain yang mempunyai kebolehan yang setara (Wright & Stone 1979; Wolins et al., 1983). Hasil dapatan bagi nilai indeks kebolehpercayaan *person* ialah 0.95 dan nilai *Cronbach Alpha Kuder-Richardson* $KR_{20}=0.97$, manakala nilai kebolehpercayaan *item* pula ialah 0.87. Nilai indeks tersebut menunjukkan masing-masing mempunyai nilai kebolehpercayaan yang diterima kuat, iaitu melebihi 0.8 (Bond & Fox, 2015). Oleh itu, perolehan daripada analisis ini menunjukkan bukti kebolehpercayaan USDAi sebagai instrumen pengukuran yang membolehkan instrumen kajian digunakan dalam penyelidikan bagi tujuan mengukur keupayaan *person* dan kesukaran *item*.

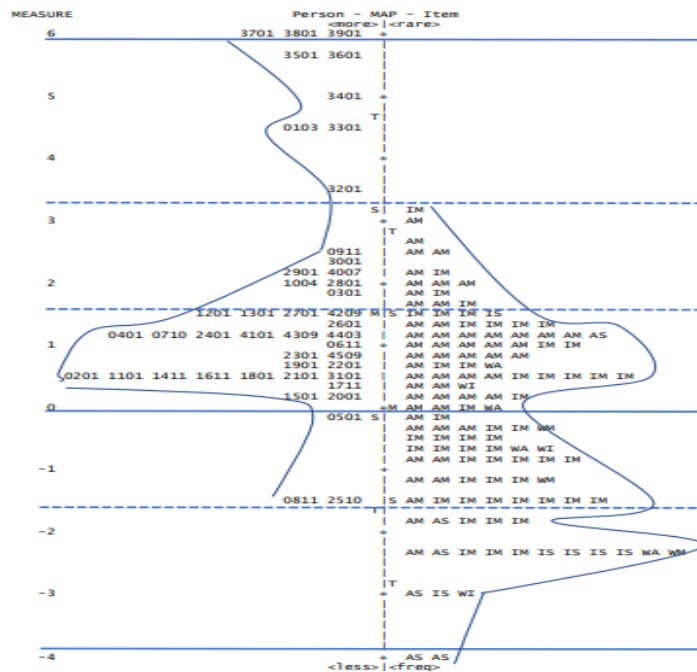
(2) Indeks Pengasingan

Person dan *item* yang diukur mengikut kumpulan dapat dikenal pasti melalui nilai pengasingan. Lebih besar nilai pengasingan atau pemisahan, semakin baik kualiti

sesuatu pengukuran. Menurut Linacre (1994), nilai pengasingan *item* bagi sesebuah instrumen sebaiknya melebihi nilai 2.0, iaitu dua tahap pengasingan. Menerusi Jadual 2 yang ditunjukkan di atas, Model Pengukuran Rasch ini dapat mengukur dua indeks pengasingan, iaitu indeks pengasingan *person* dan indeks pengasingan *item*. Indeks pengasingan *person* merujuk kepada anggaran pengasingan atau perbezaan kumpulan individu mengikut tahap kebolehan dalam pemboleh ubah atau konstruksi yang diukur. Perkara tersebut adalah bagi menunjukkan bilangan strata kebolehan responden memberikan persepsi yang dikenal pasti dalam kumpulan sampel yang diukur pada nilai 2.0. Manakala indeks pengasingan *item* menunjukkan pengasingan bagi tahap kesukaran *item* yang merujuk kepada bilangan strata kesukaran *item* pada 2 ralat piawai yang diperoleh pada set ujian yang digunakan (Wright & Stone 1979; Wolins et al., 1983; Linacre, 1994).

Perolehan indeks pengasingan USDAi menerusi Jadual 2 yang dipaparkan di atas, indeks pengasingan *person* ialah 4.33 yang menunjukkan melebihi nilai pengasingan berdasarkan standard minimum pengukuran. Nilai tersebut menunjukkan terdapatnya pengasingan bagi responden empat strata kebolehan, iaitu cemerlang, baik, sederhana dan lemah. Selain itu, Jadual 2 juga menunjukkan bahawa indeks pengasingan *item* adalah pada nilai 2.65 dan *mean infit* ialah 0.97. Dapatan ini menunjukkan bukti, bahawa *item-item* USDAi boleh dipisahkan lebih daripada tahap kesukaran, iaitu tahap *item* mudah dan tahap *item* sukar. Berikut Rajah 1 mengenai taburan pengasingan *person* dan *item*:

Rajah 1
 Taburan Pengasingan Person – Item USDAi



Dapatan indeks pengasingan person, 4.33 dan indeks pengasing item, 2.65 menerusi kajian ini menunjukkan dengan jelas pengasingan person dan item. Pengasingan person pada 4.33 menunjukkan pengasingan yang sangat baik kerana terbahagi kepada empat kelompok individu person, namun penulis mengandaikan akan lebih besar pengasingan person ini pada kajian sebenar. Hal ini kerana pemeriksaan USDAi menerusi kajian rintis ini hanya diukur ke atas 45 orang berbanding kajian sebenar nanti, iaitu penulis akan mensasarkan kepada 400 hingga 500 orang. Hal demikian juga turut berlaku kepada indeks pengasing item pada nilai 2.65 menunjukkan pengasingan yang baik kerana terbahagi kepada dua kesukaran item, iaitu item mudah dan item sukar. Namun penulis turut mengandaikan pada kajian sebenar yang akan dilakukan ke atas sampel yang lebih besar memberi perolehan pengasingan item lebih daripada dua tahap kesukaran item. Tujuan pemeriksaan instrumen selanjutnya adalah berkaitan polariti *item* atau nilai PMC dan kesesuaian *item*. Jadual 3 di bawah menunjukkan *item* fit order bagi melibat kedua-dua aspek tersebut:

Jadual 3
Item Fit Order USDAi

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PT-MEASURE		EXACT OBS%	MATCH EXP%	Item	
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP				
101	41	45	-1.44	.57	1.08	.3	9.90	4.2	A	.13	.26	90.5	90.8	B101
82	29	45	.71	.36	1.55	3.8	4.25	3.7	B	.14	.45	54.8	70.2	B082
58	16	45	2.47	.41	1.98	3.5	2.88	3.3	C	.27	.63	57.1	80.4	A058
76	12	45	3.25	.48	1.38	1.2	2.11	1.7	D	.54	.69	83.3	87.2	B076
23	25	45	1.21	.36	1.35	2.6	2.03	2.0	E	.34	.51	52.4	69.8	A023
50	22	45	1.60	.36	1.14	1.0	1.97	2.1	F	.45	.55	81.0	72.1	A050
69	23	45	1.47	.36	1.41	2.8	1.73	1.7	G	.36	.53	59.5	71.2	B069
28	19	45	2.01	.38	.97	-.1	1.67	1.6	H	.56	.59	78.6	75.5	A028
89	37	45	-.48	.43	1.47	1.7	1.63	.9	I	.14	.34	76.2	83.4	B089
114	30	45	.58	.36	1.53	3.5	1.62	1.1	J	.22	.44	47.6	70.7	B114
118	29	45	.71	.36	1.28	2.1	1.55	1.1	K	.32	.45	64.3	70.2	B118
5	42	45	-1.81	.64	1.28	.7	1.55	.8	L	.12	.23	92.9	92.9	A005
61	43	45	-2.29	.77	.89	.0	.47	-.3		.25	.20	95.2	95.3	B061
87	43	45	-2.29	.77	.83	-.1	.29	-.6		.29	.20	95.2	95.3	B087
4	44	45	-3.07	1.05	.79	.0	.13	-.7		.25	.14	97.6	97.6	A004
62	44	45	-3.07	1.05	.79	.0	.13	-.7		.25	.14	97.6	97.6	B062
124	44	45	-3.07	1.05	.79	.0	.13	-.7		.25	.14	97.6	97.6	C124
85	41	45	-1.44	.57	.74	-.6	.49	-.4	u	.36	.26	95.2	90.8	B085
88	39	45	-.90	.48	.73	-.8	.47	-.5	s	.42	.31	90.5	87.3	B088
92	40	45	-1.15	.52	.70	-.8	.42	-.5	p	.40	.29	92.9	89.1	B092
103	41	45	-1.44	.57	.67	-.8	.35	-.7	l	.39	.26	95.2	90.8	B103
13	42	45	-1.81	.64	.66	-.6	.35	-.7	k	.36	.23	92.9	92.9	A013
106	42	45	-1.81	.64	.61	-.8	.24	-.9	i	.38	.23	92.9	92.9	B106
2	43	45	-2.29	.77	.57	-.6	.13	-1.1	h	.36	.20	95.2	95.3	A002
64	43	45	-2.29	.77	.57	-.6	.13	-1.1	g	.36	.20	95.2	95.3	B064
65	43	45	-2.29	.77	.57	-.6	.13	-1.1	f	.36	.20	95.2	95.3	B065
66	43	45	-2.29	.77	.57	-.6	.13	-1.1	e	.36	.20	95.2	95.3	B066
67	43	45	-2.29	.77	.57	-.6	.13	-1.1	d	.36	.20	95.2	95.3	B067
73	43	45	-2.29	.77	.57	-.6	.13	-1.1	c	.36	.20	95.2	95.3	B073
127	43	45	-2.29	.77	.57	-.6	.13	-1.1	b	.36	.20	95.2	95.3	C127
130	43	45	-2.29	.77	.57	-.6	.13	-1.1	a	.36	.20	95.2	95.3	C130
MEAN	32.6	45.0	-.07	.49	.97	.1	.95	.0				79.7	80.1	
S.D.	8.2	.0	1.51	.23	.23	1.1	.96	.9				12.0	9.3	

(3) Polariti Item

Analisis polariti *item* merupakan langkah awal bagi memastikan *item* instrumen bergerak dalam satu arah, iaitu menepati tujuan atau objektif kajian. Hal ini untuk mengukur hubungan antara *item* dalam setiap konstruksi kajian yang ditentukan. Berdasarkan Jadual 3 di atas nilai polariti *item* bagi instrumen USDAi dalam kajian ini menunjukkan nilai *Point Measure Correlation* (PTMEA CORR) yang maksimum dan minimum dalam instrumen USDAi adalah nilai positif, iaitu tiada nilai negatif. Perkara ini dinyatakan oleh Linacre (2010), sekiranya terdapat nilai negatif pada nilai PTMEA CORR adalah menunjukkan jalinan responden pada *item* adalah bercanggah dengan konstruksi. Justeru, berdasarkan Jadual 3 di atas menunjukkan *item* adalah bergerak selari dengan *item-item* yang lain untuk mengukur konstruksi.

(4) Item Fit (Kesesuaian Item-Infit)

Pemilihan indeks *infit* MNSQ dilakukan bagi menentukan kesepadanan *item* (*item fit*), iaitu mengukur sesuatu konstruksi atau pemboleh ubah terpendam. Hal ini menurut Linacre (2002) nilai *mean square* (MNSQ) bagi setiap *item* dan responden mestilah antara 0.5 hingga 1.5. Sekiranya *item* atau responden terkeluar daripada julat tersebut memerlukan pertimbangan untuk dibaiki atau dibuang. Hal ini kerana dalam pengukuran sesuatu sifat terpendam seperti keupayaan, sesuatu instrumen hendaklah mempunyai *item-item* yang benar-benar produktif untuk mengukur sesuatu perkara. Menerusi Model Pengukuran Rasch, *item fit* adalah petunjuk kepada kualiti statistik *item* dalam menyumbang kepada pengukuran. Bond dan Fox (2015) menegaskan bahawa *mean square* (MNSQ), ialah kriteria yang diperlukan untuk menilai kesesuaian *item* tersebut.

Justeru berdasarkan Jadual 3 di atas, hasil dapatan ke atas 130 *item* telah disemak dan dapatan menunjukkan *mean infit* ialah 0.97 dan menurut Linacre (2002a) ia cukup berhasil untuk pengukuran. Terdapat tiga *item* dikesan yang mempunyai nilai *infit mean-square* (MNSQ) terkeluar melebihi julat 1.5, namun tiada *item* yang kurang daripada julat 0.5 bagi nilai tersebut. Tiga *item* yang melebihi julat 1.5 ialah *item* B082 (1.55), *item* A058 (1.98) dan B114 (1.53). Kesemua *item* ini adalah tidak sepadan atau terlalu sukar. Hal tersebut menurut Wright dan Stone (1979) dan Khoiry et al. (2019) respons yang tidak tetap tergolong dalam ketidakpadanan atau *misfit* bermaksud *item* yang melebihi nilai julat 1.5 adalah terlalu susah dan kurang nilai julat 0.5 pula merupakan *item* tersebut terlalu senang untuk responden. Kedua-dua perkara tersebut menunjukkan *item* yang dibentuk tidak begitu menguji ciri terpendam yang dikehendaki. Namun, bagi tiga *item* yang melebihi julat 1.5 menerusi pemeriksaan kajian rintis ini dikekalkan oleh penulis atas pertimbangan justifikasi tujuan kajian, iaitu mengesahkan kesukaran kosa kata imbuhan mengikut hierarki kerumitan proses derivasi berdasarkan peringkat kepelbagaian pengajian sama ada peringkat rendah, peringkat sederhana dan peringkat tinggi.

(5) Unidimensionaliti

Selain itu, ciri unidimensionaliti turut diperiksa sekiranya mempunyai ciri kesamaran dalam mengukur sesuatu perkara. Menurut Wright & Stone (1979) dan Aziz et al., (2017), ciri unidimensionaliti ternyata kritikal dalam menentukan sesuatu instrumen bagi mengukur dalam satu arah atau selari. Manakala Rohaya dan Mohd Najib (2017), unidimensionaliti merupakan pembuktian data kajian bersesuaian dengan model dan mengukur satu konstruksi yang jelas. Sehubungan itu, Aziz et al. (2017) menyatakan analisis Model Pengukuran Rasch memerlukan sekurang-kurangnya 40% dan yang baik adalah 60% bagi varians mentah (*raw variance explained by measures*) yang dijelaskan sebagai penunjuk instrumen yang baik. Sementara itu, perbezaan variasi yang tidak jelas 1 (*unexplned variance in 1st contrast*) pula hendaklah tidak melebihi had kawalan sebanyak 15% manakala bagi nilai *eigenvalue* pula tidak melebihi nilai 3.0. Perkara tersebut dinyatakan oleh Linacre (2005), jika nilai *eigenvalue* melebihi nilai 3.0 menunjukkan terdapat dimensi dalam *item*.

Jadual 4

Analisis Piawaian Varians Komponen Utama (PCA)

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)		-- Empirical --		Modeled
Total raw variance in observations	=	197.9	100.0%	100.0%
Raw variance explained by measures	=	69.9	35.3%	35.9%
Raw variance explained by persons	=	31.1	15.7%	16.0%
Raw Variance explained by items	=	38.8	19.6%	19.9%
Raw unexplained variance (total)	=	128.0	64.7%	64.1%
Unexplned variance in 1st contrast	=	16.7	8.5%	13.1%
Unexplned variance in 2nd contrast	=	15.6	7.9%	12.2%
Unexplned variance in 3rd contrast	=	8.2	4.2%	6.4%
Unexplned variance in 4th contrast	=	8.0	4.1%	6.3%
Unexplned variance in 5th contrast	=	7.4	3.7%	5.8%

Justeru, dapatan yang diperoleh menerusi Jadual 4 di atas menunjukkan unidimensionaliti dalam pembinaan instrumen USDAi, *raw variance explained by measures* menunjukkan 35.3% berbanding dengan Model Pengukuran Rasch, iaitu 35.9% adalah menghampiri dengan batasan 40%. Namun menurut Fisher (2007), nilai keperluan bagi instrumen adalah sekurang-kurangnya 20% dan ia menunjukkan *raw variance explained by measures* bagi USDAi telah tercapai dan boleh diterima. Tahap *noise* yang diukur pula adalah sebanyak 8.5% (*unexplned variance in 1st contrast*) adalah boleh diterima memandangkan ia jauh daripada nilai terkawal tidak melebihi 10% atau tidak melebihi maksimum had kawal 15%. Keadaan ini menunjukkan tidak wujud dimensi kedua. Manakala nilai *eigenvalue* sebanyak 16.7 menunjukkan nilai yang melebihi nilai 3.0 menunjukkan terdapat dimensi dalam *item* (Linacre, 2005). Hal tersebut kerana *item* USDAi yang dibentuk sebanyak 130 *item* diuji ke atas 45 orang sampel yang menunjukkan bilangan sampel tersebut adalah terlalu kecil berbanding *item* yang diuji. Dalam perkara tersebut, penulis akan menjangkakan kajian sebenar yang melibatkan bilangan yang lebih besar akan meningkatkan pembinaan nilai unidimensionaliti seperti disarankan oleh Linacre (2005). Namun secara keseluruhan,

item bagi kesemua *item* dan konstruk USDAi yang diuji menerusi model dikotomi ini adalah memenuhi optimum andaian unidimensionaliti dan kesahan konstruksi berdasarkan Model Pengukuran Rasch.

Perbincangan

Kajian ini dapat membina sebuah instrumen Ujian Struktur Dalam Kata Imbuan (USDAi) yang sah dan boleh dipercayai bagi menguji penguasaan pelajar bahasa Arab. Hal ini dibuktikan USDAi dapat mengukur penguasaan pelajar bahasa Arab sebagai bahasa kedua atau asing di Malaysia dan mengesahkan kesukaran kosa kata imbuan mengikut lapisan kerumitan proses derivasi. Pendekatan model pembangunan instrumen oleh McIntire dan Miller (2007) dan Model Pengukuran Rasch oleh George (1980) turut membantu memeriksa *item* dan instrumen USDAi yang dibina dengan lebih teliti dan jitu. Hal tersebut kerana Model Pengukuran Rasch yang mengukur *item* dan konstruksi membuktikan instrumen yang dibina mempunyai indeks kebolehpercayaan yang tinggi.

Perkara tersebut ditegaskan kerana kajian-kajian lepas yang dilakukan khususnya di Malaysia seperti kajian Abu Bakar (2008), Jalil (2012), Rasad dan Mahamod (2013), Norazwa (2014), Baharudin et al. (2017), Asbullah et al. (2018), Maskor et al. (2018) dan Tuan Omar et al. (2018), kebanyakannya tertumpu kepada penguasaan dan permasalahan morfologi terutamanya aspek kata mahupun kosa kata Arab. Manakala pengesanan permasalahan yang ditemui, menggunakan instrumen yang tidak bersandarkan teori yang kukuh dalam pembinaan sesuatu instrumen bagi mengukur dan menguji permasalahan aspek kata dalam bidang morfologi Arab.

Signifikan permasalahan kajian berhubung aspek kata dan strukturnya menerusi morfologi ini menyebabkan keupayaan para pelajar terganggu dan terhalang untuk menguasai sesebuah bahasa terutamanya bahasa Arab. Hal ini disahkan menerusi pengujian USDAi yang dilakukan, para pelajar sering melakukan kesilapan menentukan huruf akar daripada struktur terbitan yang kompleks. Implikasinya, para pelajar menghadapi kesukaran untuk memahami makna asas perkataan dan luas kata yang sebenar lebih-lebih lagi menggunakan sesuatu perkataan dalam konteks ayat mahupun pertuturan mereka.

Justeru menerusi kajian ini, penulis melaporkan hasil dapatan yang baharu berkaitan bidang morfologi Arab, iaitu penghasilan sebuah instrumen USDAi yang berpandukan kepada teori struktur dalaman kata (SDK) dan teori paradigmatis serta konsep derivasi bahasa Arab bagi menentukan penguasaan pelajar terhadap struktur kata imbuan Arab. Hasil pemeriksaan terhadap kesahan dan kebolehpercayaan USDAi yang dibina berdasarkan teori dan konsep struktur dalaman kata (SDK), mendapati item-item USDAi benar-benar produktif untuk mengukur konstruk kata akar, imbuan dan *wazan* kata kerja imbuan bahasa Arab.

Hasil pengujian yang dilakukan menerusi kajian rintis, mengesahkan USDAi merupakan sebuah skala atau alat parameter yang mampu mengukur penguasaan struktur dalaman kata (SDK) Arab di samping mengesahkan hierarki kesukaran *item*

secara objektif dan statistik. Perkara tersebut dibuktikan daripada analisis yang dilakukan menerusi USDAi, bahawa terdapat jenis kata yang sukar dikuasai oleh para pelajar. Berikut jenis kata yang dikenal pasti:

- (1) Jenis kata kerja *al-fi' al-madiy* seperti آمَن dan أَحْيَا, kata kerja *fi' al-'amr* أَجْرٌ dan أَجَبٌ dan kata nama ism al-fa' il مُجْتَرِيٌّ .
- (2) Kata kerja *al-fi' al-madiy* bagi آمَن , أَحْيَا , أَجْرٌ , أَجَبٌ dan مُجْتَرِيٌّ merupakan jenis perkataan *mahmuz al-fa'* (مهموز الفاء),
- (3) Perkataan أَجْرٌ dan أَحْيَا merupakan jenis kata *'illah* iaitu satu huruf vokal atau *'illah* pada pertengahan perkataan atau dua huruf vokal di akhir perkataan dan perkataan أَجَبٌ bagi padanan *wazan* أَفْعِلٌ jenis perkataan *al-muda'af* (المضاعف), iaitu huruf berganda (*al-Shaddah*).

Signifikan hierarki kesukaran tersebut adalah merujuk kepada teori teori SDK, *paradigmatis* dan konsep derivasi Arab bagi kajian ini. Hal ini disebabkan proses pembentukan perkataan (*word formation*) daripada kata akar dan huruf tambahan (imbuhan) yang berpandukan kepada *wazan* kata kerja imbuhan bahasa Arab (imbangan pola) sama ada melalui pengimbuhan (*mazid*), pengguguran (*hazaf*), pengalihan (*i'āl*), penggantian (*ibdāl*) mahupun penyerapan (*idghām*) merupakan penggolongan perkataan dalam bahasa Arab. Perkara tersebut dinyatakan kerana proses pembentukan perkataan Arab adalah berdasarkan sistem turunan morfem derivasi yang berdasarkan pecahan kata akar kemudian menerbitkan pelbagai bentuk derivasi yang berbeza. Selain itu, proses pembentukan kata tersebut turut mencipta leksem baru bagi memperluaskan leksikon bahasa (kosa kata bahasa).

Berdasarkan teori SDK, *paradigmatis* dan konsep derivasi tersebut, perubahan pemeringkatan konstruksi dan lapisan hierarki semasa proses pengimbuhan (*mazid*), pengguguran (*hazaf*), pengalihan (*i'āl*), penggantian (*ibdāl*) mahupun penyerapan (*idghām*) merupakan faktor kesukaran pelajar dalam menguasai struktur dalaman kata (SDK) imbuhan bahasa Arab. Selain itu, pertimbangan kesahihan dapatan daripada instrumen yang dijalankan memberi pembuktian empirikal dan benar, iaitu data mentah menerusi ujian rintis ini dianalisis menggunakan pendekatan analisis Model Pengukuran Rasch yang berupaya menganalisis ketidaksepadanan atau ketindanan *item* dalam setiap konstruksi. Instrumen yang dihasilkan juga menepati keperluan dan tujuan kajian berikutan konstruksi atau dimensi yang dibina dapat menjelaskan dapatan kajian secara bersasar. Manakala, indeks kebolehppercayaan *item* dan responden USDAi yang diperolehi turut menunjukkan keabsahan dan keupayaan untuk mengukur penguasaan struktur dalaman kata (SDK) imbuhan, di samping menunjukkan hierarki kesukaran *item* secara objektif dalam kalangan pelajar bahasa Arab sebagai bahasa kedua atau bahasa asing.

Secara rumusan pengesanan instrumen USDAi menerusi makalah ini, keseluruhan 130 *item* yang dibentuk mampu mengukur penguasaan struktur dalaman kata (SDK) imbuhan bahasa Arab. Data kajian yang diperolehi turut menepati model pengukuran yang membuktikan USDAi merupakan instrumen yang mempunyai ciri

kesahan konstruksi dan darjah kebolehppercayaan yang tinggi untuk digunakan dalam penyelidikan sains sosial. Manakala, pengaplikasian Model Pengukuran Rasch dalam instrumen USDAi yang dibina adalah tepat kerana kesahan dan kebolehppercayaan dianggap bermanfaat berdasarkan keupayaan model. Sehubungan itu, data yang diperoleh dapat menggambarkan pembinaan *item-item* adalah sah dan menyampaikan tujuan yang jelas mengenai pembinaan instrumen bersesuaian dengan jangkaan teori. Justeru, penghasilan instrumen USDAi bagi penguasaan morfologi pelajar bahasa Arab berdasarkan pengaplikasian Model Pengukuran Rasch ini boleh dilaksanakan bagi menguji kualiti penguasaan pelajar yang mempelajari bahasa Arab di Malaysia

Kesimpulan

Kajian ini bertujuan melaporkan kesahihan konstruksi dan kebolehppercayaan instrumen rintis bagi mengukur kesukaran struktur dalaman kata (SDK) imbuhan bahasa Arab. Instrumen yang dibina membuktikan bahawa kesahan pengukuran dan kebolehppercayaan instrumen USDAi bagi penguasaan pelajar bahasa Arab dapat dicapai setelah melakukan pemeriksaan yang mematuhi piawaian kesahan yang ditetapkan bahkan menepati reka bentuk instrumen dalam kajian ini. Penemuan utama kajian menerusi hasil atau *outcome* kajian, instrumen USDAi yang dihasilkan adalah bersifat sah dan jitu yang menjadi skala pengukur (*measure*) atau satu parameter (*ruler*). Hal ini kerana lazimnya skala pengukur ini hanya digunakan dalam bidang sains tulen seperti fizik, kimia dan biologi. Namun, menerusi kajian ini mampu menghasilkan sebuah skala pengukur bagi kegunaan penyelidikan bidang sains sosial seperti bahasa Arab. Skala pengukur ini bertujuan para pelajar dapat memberi maklum balas mengenai penguasaan dan keupayaan pelajar terhadap kata akar, huruf tambahan (imbuhan) dan *wazan* kata kerja imbuhan bahasa Arab berdasarkan pengetahuan sedia ada selari mengikut kognitif mereka. Selain itu, penemuan jenis kata *mahmuzat*, *‘illah* dan *muda‘af* yang melalui proses derivasi adalah paling sukar dikuasai oleh para pelajar bahasa Arab. Isu tersebut memberi implikasi kepada penguasaan kosa kata dan perkembangan kemahiran bahasa. Hal ini kerana penguasaan struktur dalaman kata (SDK) bagi sesuatu perkataan Arab seperti mengetahui bentuk-bentuk kalimah dan proses perubahan bentuk kata imbuhan merupakan penguasaan asas yang sewajarnya dikuasai terlebih dahulu berbanding penguasaan makna perkataan dan penggunaannya dalam ayat. Sekiranya perkara ini tidak diatasi menyebabkan hilangnya keharmonian dan keindahan sistem sesuatu bahasa. Kajian ini adalah terbatas kepada pengesanan permasalahan berkaitan struktur dalaman kata (SDK) menerusi pembinaan instrumen yang dibina. Di samping itu, kajian ini dapat melaporkan penguasaan pelajar berkaitan penguasaan kata akar, imbuhan dan *wazan* kata kerja imbuhan bahasa Arab. Walau bagaimanapun, dapatan kajian ini bukan muktamad kepada penyelesaian isu-isu yang berkaitan kelemahan pembelajaran morfologi bahasa Arab terutama penguasaan struktur dalaman kata (SDK). Justeru, dicadangkan agar para penyelidik akan datang menimbulkan persoalan baharu berkaitan dengan isu berkaitan SDK seperti pembinaan modul SDK, implimentasi modul SDK dan tinjauan keberkesanan modul SDK menerusi

kajian eksperimen. Usaha tersebut adalah penting dalam memastikan bahasa Arab kekal relevan dipelajari dan dikuasai sebagai bahasa asing di Malaysia.

Penghargaan

Penulis merakamkan ucapan terima kasih kepada Bahagian Biasiswa dan Pembiayaan, Kementerian Pelajaran Malaysia (KPM) atas pemberian sokongan dana di bawah program Hadiah Latihan Persekutuan (HLP) Tahun 2018.

Rujukan

- Abu Bakar, K. (2008). *Rasch calibration of Arabic reading diagnostic test items*. International Islamic University.
- Abu Bakar, K., & Hassan, S. H. (2011). Applying the Rasch model in educational researches. In F. A. Rahman, F. Piee@Shafiee, & H. Elias (Eds.), *Teachers' learning curriculum innovations and knowledge applications* (pp. 110–127). UPM Press.
- Alderson, J. C. (1998). Developments in language testing and assessment, with specific reference to Information Technology. *Forum for Modern Language Studies*, 2, 195–206.
- Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (Part I). *Language Teaching*, 34(4), 213–236. <https://doi.org/10.1017/S0261444800014464>
- Allan, A. (1992). Development and validation of a scale to measure test-wiseness in EFL/ESL reading test-takers. *Language Testing*, 9, 101–122.
- Al-Hafiz Abd al-Rahim, A.-S. (2006). *Mabadi' Ta'lim al-Lughah al-Arabiyyah li Ghair al-Natiqin biha* (1st ed.). Jidaran lil kitab al-Alami
- Al-Najem, S. R. (2007). Inheritance-based approach to Arabic verbal root-and-pattern morphology. In A. Souidi, A. van den Bosch & G. Neumann (Eds.), *Arabic computational morphology* (pp. 67–88). Springer. https://doi.org/10.1007/978-1-4020-6046-5_5
- Asbulah, L. H., Lubis, M. A., & Aladdin, A. (2018). Kesahan dan kebolehpercayaan instrumen pengetahuan kolokasi Bahasa Arab IPT (I-KAC IPT) menggunakan model pengukuran Rasch. *Asean Comparative Education Research Journal On Islam And Civilization (ACER-J)*. EISSN2600-769X, 2(1), 97–106.
- Azami, Z., Afida, A., Mohd Zaidi, O., & Norazreen, A. A. (2019). Pengukuran terhadap penilaian psikometri. In M. N. Zulkifli, O. Siti Aminah & A. Norhana (Eds.), *Pengukuran Rasch dalam penilaian hasil pembelajaran pendidikan kejuruteraan* (Pertama, pp. 102–189). Universiti Kebangsaan Malaysia.
- Aziz, A. A., Masodi, M. S., & Zaharim, A. (2017). *Asas model pengukuran Rasch: Pembentukan skala & struktur pengukuran* (ke3 ed.). Universiti Kebangsaan Malaysia.
- Baharudin, H., Ismail, Z., & Nik Yusoff, N. M. R. (2017). Pembinaan ujian tahap saiz kosa kata Arab the construction exam level size of Arabic vocabulary. *Jurnal Pendidikan*

- Malaysia*, 42(1), 13–19.
- Bloomfield, L. (1970). *Language*. George Allen and Unwin Ltd.
- Bond, T. G., & Fox, M. C. (2015). *Applying The Rasch model fundamental measurement in the human sciences* (3rd ed.). Routledge Taylor & Francis Group.
- Booij, G. E. (2015). The structure of words. In *The Oxford Handbook of the Word* (Issue December, pp. 1–19). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199641604.013.002>
- Buckwalter, T., & Parkinson, D. (2011). *A frequency dictionary of Arabic: Core vocabulary for learners*. Routledge.
- Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics*, 20, 147–161.
- Cumming, A. (1996). Introduction: The concept of validation in language testing. In A. Cumming & R. Berwick (Eds.), *Validation in language testing*. Multilingual Matters.
- Fisher, W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions*, 21(1), 1095.
- George, R. (1980). *Probabilistic models for some intelligence and attainment tests*. The University of Chicago Press.
- Ghizawiy, B. (2008). *Mu'jam Tasrif al-'A^fal al-Madrasiy*. Dar al-^oIlmi al-Malayiyin.
- Gregory, C., Charlotte, A., & Silki, L. (2013). *Measurement in education in the United States*. Oxford University Press.
- Hassan Basri, A. M. D. (2002). *Language testing: The construction and validation*. University of Malaya Press.
- Khoiry, M. A., Nik Ibrahim, N. L., Osman, M. H., & Hamzah, N. (2019). Isu Kesahan dan Kebolehppercayaan Penilaian Pelajar. In Z. M. Nopiah, S. A. Osman, & N. Arsad (Eds.), *Pengukuran Rasch dalam penilaian hasil pembelajaran pendidikan kejuruteraan* (pp. 190–250). Universiti Kebangsaan Malaysia.
- Kusaeri, S. (2012). Pengukuran dan penilaian pendidikan. *Buku Graha Ilmu*, 1(1), 1–10. info@grahailmu.co.id
- Linacre, J. M. (1994). *Sample size and item calibration stability*. *Rasch Measurement Transactions*, 7(4), 328.
- Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Linacre, J. M. (2005). *A user's guide and program manual to Winsteps: Rasch Model computer program* (3.6.1.11). MESA Press.
- Linacre, J. M. (2010). Predicting responses from rasch measures. *Journal of Applied Measurement*, 11(1), 1–10.
- Linacre, J. M. (2021). *A user's guide to WINSTEPS: Rasch-model computer program* (5.0.2.0). MESA Press.
- Maeda, J., & Naoumi, E. (2005). Vocabulary Development Using a Questionnaire. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT2004 Conference Proceedings* (pp. 732-742). JALT5t.
- Maher Sya^oban, A. al-B. (2011). *Ta^olim al-Mufradat al-Lughawiyah*. Dar al-Musayyarah Li al-Nasyr wa al-Tauzi^o

- Maskor, Z. M., Baharudin, H., & Lubis, M. A. (2018). Measurement validity and reliability of the productive vocabulary knowledge instrument for Arabic learners in Malaysian secondary schools. *Advanced Science Letters*, 24(5), 3423-3426.
- McIntire, S. A., & Miller, L. A. (2007). *Foundations of psychological testing: A practical approach* (2nd ed.). SAGE Publications.
- Jalil, A. M. F. (2012). *Penguasaan pelajar STAM terhadap imbuhan kata Bahasa Arab*. Universiti Malaya.
- Muhammad Ma'sum, A. (1965). *al-Amthilah al-Tasrifiiyyah Lil Madaris al-Salafiiyyah al-Syafi'eyyah*. Maktabah wa Matba'ah Salim Nabhan.
- Muhammad Qadur, A. (1999). *Madahkhil 'Ila Fiqh al-Lughah al-'Arabiyyah*. Dar al-Fikr.
- Norazwa, A. R. (2014). *Penguasaan pelajar Sarjana Pengajian Islam terhadap Al-'Afr' al Al-Mazidah dalam morfologi Arab*. Universiti Kebangsaan Malaysia.
- Perneger, T. V., Courvoisier, D. S., Hudelson, P. M., & Gayet-Ageron, A. (2014). Sample size for pre-tests of questionnaires. *Quality of Life Research*, 24(1), 147–151. <https://doi.org/10.1007/s11136-014-0752-2>
- Rabiah Tul Adawiyah, M. S., Biase, B. Di, & Wan Nur Madiha, R. (2020). The acquisition of english grammar among malayenglish bilingual primary school children. *GEMA Online Journal of Language Studies*, 20(4), 166–185. <https://doi.org/10.17576/gema-2020-2004-10>
- Rahman, M. S. (2016). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language “Testing and Assessment” research: A literature review. *Journal of Education and Learning*, 6(1), 102. <https://doi.org/10.5539/jel.v6n1p102>
- Rasad, F., & Mahamod, Z. (2013). Analisis kesalahan morfologi dan sintaksis dalam penulisan tesis cemerlang: Satu kajian kes. In Z. Mahamod, A. Ahmad, A. H. Zakaria, N. N. F., Nik Harmi (Eds.), *Prosiding Seminar Pascasiswazah Pendidikan Bahasa Melayu & Kesusasteraan Melayu* (pp. 311–335). Fakulti Pendidikan, Universiti Kebangsaan Malaysia.
- Rohaya, T., & Mohd Najib, A. G. (2017). Pembinaan dan pengesahan instrumen bagi mengukur tahap literasi pentaksiran guru sekolah menengah di Malaysia. *Noise and Health*, 19(89), 165–173. https://doi.org/10.4103/nah.NAH_106_16
- Rosni, S. (2009). *Pendekatan pembelajaran kemahiran bahasa Arab untuk pelajar bukan penutur jati*. Universiti Sains Islam Malaysia.
- Song, F. (2002). A comparative study on the productive and academic vocabulary knowledge of Japanese and Chinese university students: Focusing on the students' performances in the two vocabulary tests. *Journal of International Development and Cooperation*, 9(1), 113–127.
- Steven, J. R. (2008). Language testing in Asia: Evolution, innovation, and policy challenges. *Language Testing*, 25(1), 5–13. <https://doi.org/10.1177/0265532207083741>
- Tuan Omar, T. N. A., Zailani, S., Abu Bakar, K., Zainal, H., Nawawi, Z., & Nik Farhan, M. (2018). Mengesan kelemahan morfologi Arab melalui ujian diagnostik morfologi Arab (UDMA) di SMK Agama Miri. *Prosiding Nadwah Ulama Nusantara (Nun) Viii*

2018 Sinergi Ulama Dalam Pemerakaan Ummah, 803–812.

- Tucker, M. A. (2011). *The morphosyntax of the Arabic verb: Toward a unified syntax-prosody*. <https://escholarship.org/uc/item/0wx0s7qw>
- Ullah, U., Ramanair, J., & Rethinasamy, S. (2021). Development and validation of willingness to communicate, language use, and motivation questionnaires. *Issue in Language Studies*, 10(1), 20–36.
- Wolins, L., Wright, B. D., & Masters, G. N. (1983). Rating scale analysis: Rasch measurement. *Journal of the American Statistical Association*, 78(382), 497. <https://doi.org/10.2307/2288670>
- Wright & Stone (1979). *Test design*. MESA Press, 57-70 [https://doi.org/10.1016/s1567-8032\(00\)80032-4](https://doi.org/10.1016/s1567-8032(00)80032-4)